4-26-2019

Genetic Diversity in the Leucoagaricus Cultivar Fungus of Trachymyrmex septentrionalis Fungus-Growing Ants

Lee Deininger
lee.deininger@uconn.edu

Genetic Diversity in the *Leucoagaricus* Cultivar Fungus of *Trachymyrmex septentrionalis* Fungus-Growing Ants

Lee Deininger

B.S., University of Connecticut, 2016

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

At the

University of Connecticut

2019

APPROVAL PAGE

Masters of Science

Genetic Diversity in the *Leucoagaricus* Cultivar Fungus of *Trachymyrmex septentrionalis* Fungus-Growing Ants

Presented by

Lee Deininger, B.S.

Major Advisor_____
Jonathan Klassen

Associate Advisor_____
J. Peter Gogarten

Associate Advisor_____
Joerg Graf

Associate Advisor_____
David Benson

University of Connecticut

2019

Acknowledgments

I would like to thank my major advisor, Jonathan Klassen, for the infinite wisdom, patience, and guidance he gave me during my time in the lab and for shaping me into the scientist I am today. Thanks to everyone in the Klassen lab, past and present, for their help and friendship; and for creating a fantastic environment for injury and growth. Thanks to the UCONN MARS facility and to CGI for the sequencing services. Thanks to my committee: Joerg Graf, J. Peter Gogarten, and David Benson for taking the time to question and encourage this project to come to fruition. Thanks to my friends and family for supporting me throughout this project. Thanks to all fungi for mystifying to us and for keeping the planet alive. Special thanks to Stephanie (not a fungus), whom all of this was meant to impress.

Table of Contents

1 – Introduction

1.1 – The Kingdom Fungi

The eukaryotic Kingdom Fungi contains mushrooms, yeasts, molds, lichens, mycorrhizae, rusts, and smuts, organisms with extremely diverse life histories. Their unique competitive strategies, ecological positions, and methods of reproduction have made them extremely biodiverse. This remarkable diversity has greatly benefited ecosystems across the planet because fungi are essential symbionts of many organisms (Blackwell, 2011; Stajich et al., 1996). Our understanding of the diversity and complexity of fungal life strategies, however, is still underdeveloped.

Recent studies using high-throughput sequencing methods suggest that fungi have at least 6 times more species than plants, totaling an estimated 5.1 million fungal species. Using molecular methods to discover and stably classify fungal taxa led to 60,000 new fungal species being described between 1943 and 2008. This has been greatly facilitated by the adoption of a single conserved and highly informative gene for phylogenetic inferences: the nuclear ribosomal internal transcribed spacer (ITS) gene. This has resolved early diverging fungal lineages and allowed for the identification of a broad range of fungi (Schoch et al., 2012). As such, it has become the gold standard for fungal phylogenetics.

Fungal diversity includes a broad array of life strategies, the complexity of which challenges our explanatory models and terminology (Burnett, 1975). These life strategies are distinguished by the occurrence, sequence of, and products of sexual and asexual reproduction, and the occurrence of hyphal fusions, which have profound effects on fungal biology. The number of fungal nuclei and their genetic composition can vary greatly between life cycle and

sexual stages. To understand the diversity of fungal genetics, one must establish the foundations for the typical life cycle of fungi and then identify where these can differ.

1.2 - Genetic diversity - patterns and processes

Fungal genetics are best understood in fungi that reproduce sexually. These include most described fungi, apart from the Fungi Imperfeci (Deuteromycota) and Mycelia sterila (Burnett, 1975). Other groups that do not reproduce sexually constitute some 25,000 species and include many Basidiomycota and Ascomycota anamorphs (whose morphologies differ from their teleomorph forms and who reproduce via conidia or asexual spores), and other fungi that do not produce any spores, sexual or otherwise.

Sexual reproduction in fungi contains three basic stages: *plasmogamy,* the fusion of cytoplasm between conjugate cells; *karyogamy*, the fusion of conjugant nuclei; and *meiosis*, the reductive division of the fused nuclei following karyogamy. These stages have been used to describe many disparate sexual life cycles in fungi. The most common sexual fungal life cycle – or, at least the best understood – is one that shifts from *monokaryotic*, containing a single haploid set of parental genes, to *dikaryotic*, containing two haploid gene sets. This occurs during the transition from mycelia to mushroom production in Basidiomycetes and is referred to as a *haploid dikaryotic* lifecycle. Here, a haploid mycelium derived from the germination of a meisospore persists until plasmogamy occurs, which is marked by a shift to unrestricted and independent mycelial growth and constitutes the longest-lived dikaryotic phase of this fungal life cycle. While dikaryotic, potentially conjugant nuclei persist close to each other in a fused hypha and divide synchronously for a variable amount of time depending on the fungal species. As these hyphae divide, they become briefly monokaryotic. To maintain a stable dikaryotic state, one nucleus may enter a special compartment at nuclear division referred to as a *clamp connection*. This intercompartmental fusion forms linkages between hyphal cells that equally distribute each conjugate nucleus between these cells as their nuclei divide. Fungi transport,

fuse, and recombine their nuclei using many mechanisms, in which cell-to-cell karyotic conditions may vary (Burnett, 1975). In plasmogamy that leads to karyogamy, fungal hyphae fuse to form a continuous cytoplasm. In normal conditions, karyogamy does not immediately occur between the resulting dikaryotic nuclei but instead occurs later on after the formation of specialized cells such as *basidia* or *asci*, the reproductive structures of basidiomycetes and ascomycetes, respectively. Sometimes, karyogamy occurs within normal vegetative hyphae, a process called *somatogamy*. Here, conjugant nuclei migrate to the sites of karyogamy and divide synchronously within a continuous cytoplasm.

When many heterogeneous nuclei coexist in a common cytoplasm for extended periods of time, these cells are referred to as *heterokaryons* (Burnett, 2003). Heterokaryons can be formed from fungal strains with genetically different nuclei that may or may not share the same mating type. However, only heterokaryons with nuclei having the same mating type are compatible for karyogamy. In nature, heterokaryons occur spontaneously, although their stability is highly variable leading to a reversion to cells with homogenous nuclei or nucleus reshuffling between cells to form new heterokaryons. Heterokaryosis therefore leads to genetic plasticity within a mycelium and can enable recombination and segregation of component nuclei that are governed by evolutionary selection.

Heterokaryosis is often associated with *parasexuality*, a distinct life strategy that often occurs in place of, or in addition to, sexual reproduction (Burnet, 1975, 2003). Like sexual reproduction, parasexuality also involves three stages: i) hyphal fusion and plasmogamy of two mycelia containing genetically different nuclei; ii) karyogamy in *vegetative* hyphae (*not* in specialized cells such as basidia or asci); and iii) somatic recombination, which may or may not be followed by non-meiotic reductive divisions of the diploid nucleus to restore haploidy. This life strategy is often inferred from extreme genetic diversity and variability in strictly asexual fungi,

e.g., as in the plant pathogen *Magnaporthe grisea*, the causative agent of wheat blast (Zeigler et al., 1996).

Parasexuality is not the only way by which single fungal cells can contain genetically diverse nuclei. For example, *polyploidy* occurs when three or more sets of heterogeneous parental genes are contained within a *single* hybridized nucleus (Burnett, 2003). Polyploidy has been linked to evolutionary innovation following whole-genome duplications in fungi such as *Saccharomyces cerevisiae,* an ancient polyploid that has reverted to a diploid state (Albertin & Marullo, 2012). This process (*diploidization*) can sometimes lead to the loss of entire chromosomes, genome rearrangement, and sequence divergence, together resulting in differentiation of the duplicated chromosomes (Burnett, 2003). Polyploidy can therefore generate new sets of chromosomes that originate from a single species or via hybridization between closely related species. However, hybridization events via polyploidy in Ascomycota and Basidiomycota experience strong intraspecific antagonism and are repressed by widespread mechanisms that regulate hyphal fusion between species.

In all of the above cases, hyphal fusion is followed by the completion of sexual reproduction. This contrasts with *asexual* life cycles that do not complete sexual reproduction. Because this definition is based on the absence of a phase, it is somewhat artificial but remains convenient to describe the situation of many fungi (Burnet, 1975). Asexual fungi are essentially clonal in nature and reproduce non-sexually via mitotically-produced spores, conidia, and hyphal fragments: asexual propagules perpetuate the original clone. If clonal mycelia grow close together, they can form *anastomoses*, linkages between multiple fungal cells that create a continuous mycelium. This also describes anamorphic forms of true fungi that occur separate from their respective teleomorphs, a *pleomorphism* where a single fungus has distinct forms depending on whether it is in a sexual or asexual phase (Cole & Kendrick, 1981). Fungi

therefore exhibit diverse modes of sexual and asexual reproduction that can vary in different fungal life cycle stages.

1.3 - Fungi as adapted mutualists

Fungi participate in diverse associations with other species (Cooke, 1977). Some are species-specific, reproductively dependent on a host, and feature asexual fungi with low genetic diversity. Others include diffuse associations with multiple free-living fungi. Often, these symbioses dictate the life cycles and genetic condition of the fungal partner. Unravelling the mechanisms that influence these dynamics will improve our knowledge of how symbionts interact and diversify in mutualisms. Some plant and insect symbioses feature associations between fungi and hosts that are immensely beneficial to both partners. For such long-lasting and interdependent symbioses, understanding the phylogenetic diversity of each member can reveal how symbioses are mediated by patterns of phylogenetic diversity. However, how fungi reproduce and thereby influence the dynamics of such symbiotic systems often remains unknown.

One example of such symbioses is the association between arbuscular mycorrhizal fungi (AMF) and certain plants (Hass & Kerp, 1994). These fungi benefit plants by increasing nutrient acquisition beyond what their roots could access alone. In exchange, plant hosts provide AMF with stable access to other nutrients such that both organisms mutually benefit from the interaction. AMF form symbioses with most known plant species dating back to approximately 400 million years ago and influence plant species diversity and productivity (Croll & Sanders, 2009; Krishna 2005). AMF were originally thought to be entirely clonal, reproducing only via asexual spores, even though this would contradict evolutionary theory, e.g., totally clonal organisms should accumulate deleterious mutations that lead to eventual extinction (Muller, 1964; Zeigler et al., 1996) . However, recent studies have reported evidence suggesting recombination in AMF, leading to genetically diverse nuclei within a single clonal network of

fungal cells (Croll & Sanders, 2009; Zeigler et al., 1996). The genetic communities that exist within AMF are therefore more consistent with long-term heterokaryosis, which explains why AMF might have different ribosomal genes despite never sexually reproducing. AMF might therefore be better described as conglomerates of individuals or genetic populations, despite being morphologically singular entities (Krishna, 2005). These results suggest that AMF have evaded extinction despite being asexual by using heterokaryosis to maintain genetic diversity. How these genetic mechanisms influence interactions between AMF and their plant hosts remains largely unknown.

The diversity of fungal symbionts is not always governed by genetic mechanisms, such as when a flexible set of fungal symbionts associate with a single host species. One example of such a symbiosis is that between bark beetles and their symbiotic fungi (Six, 2007). In this symbiosis, beetles transmit fungi between various tree species as they penetrate their bark. The beetle primarily benefits by feeding on the fungi as they grow within the tree, and the fungus benefits by being transported to new suitable habitats. Certain species of bark beetle associate with up to three different fungal symbionts, each specialized to different regions of the beetle's geographic range (Sperling, 2011). Which fungi predominate depends on ecological mechanisms such as the geographic location of their host and the degree of polyculture and/or host switching. Having a diverse set of fungal partners allows these beetles to successfully colonize trees throughout their range.

Symbiotic fungal diversity can also be influenced simultaneously by both genetic and ecological factors. For example, fungus-growing termites (Macrotermitinae) have coevolved with a single fungal genus *Termitomyces,* although horizontal symbiont transfer occurs in most termite genera (Korb & Aanen, 2002). Like other agricultural mutualisms, *Termitomyces* benefits its host by providing nutrition and the termites provide a suitable habitat for the fungus. Each termite colony contains a single variant of *Termitomyces* despite using genetically variable,

horizontally transferred spores as an inoculum (Kerstes et al., 2009). The clonally-related mycelia of these cultures fuse within the colony (Korb & Aanen, 2002), which increases their efficiency of spore production. In this way, fungal diversity is influenced by ecological mechanisms (host cultivation of a single, horizontally-acquired fungal variant) and by genetic mechanisms enacted by the fungal symbiont (hyphal fusion leading to heterokaryosis).

Associations between fungi and host organisms are as diverse as they are varied, and include different degrees of intimacy, permanence, and nutritional and reproductive interdependence (Cooke, 1977). The importance of specificity within these associations can vary, but the genome dynamics of each fungal symbiont fungi can substantially influence the functionality of microbial ecosystems (Wilkinson et al., 2010). Understanding fine-scale diversity in fungal communities is therefore important to understand the contribution of such diversity to ecosystem processes.

1.4 – The fungus-growing ant mutualism

One well-studied symbiotic partnership that is a model for symbiotic specificity and persistence is that which occurs in ants that have evolved agriculture – the attini. The attini are a monophyletic tribe of *Myrmicinae* that obligatorily depend upon the cultivation of fungi within fungal gardens as their primary source of nutrition (Schultz & Brady, 2008). Many methods of fungal cultivation, or agricultural types, have arisen throughout the evolution and diversification of this ant tribe, although most cultivate fungi from a single family of fungi: *Agaricacae* (Mueller et al., 1998). These agricultural types include coral agriculture, basal agriculture, yeast agriculture, domesticated agriculture, and leaf-cutter agriculture, which are distinguished by the type of the fungi that they cultivate and the fidelity with which the ants maintain their cultivars. Most of the more basal types of fungus-growing ant host fungi that can live without their ant hosts (Mueller et al., 1998). This includes coral fungus agriculture, which is conducted by *Apterostigma* ants (paleoattini) who cultivate a fungus from the *Pterulaceae* instead of the

Agaricacae cultivar grown by all other types of ant agriculture (Mueller et al., 1998). In yeast agriculture, *Cyphomyrmex* ants rear yeast morphotypes the *Agaricacae* cultivar fungus. This is the only agriculture type that involves non-hyphal fungi. Another primitive agricultural type is basal agriculture, conducted by paleo- and neoattini genera such as *Myrmicocrypta*, *Cyphomyrmex*, and *Apterostigma*. Ants in this agricultural type rear a range of fungi capable of both living outside of the symbiosis and sexual reproduction (Mueller et al., 1998) and commonly engage in cultivar swapping (Mikheyev et al., 2010).

Leaf-cutter agriculture, the most evolutionarily advanced system, is conducted by *Atta* and *Acromyrmex* ants. These ants have adapted highly specialized castes and morphological adaptations to suit the symbiosis, and associate with a single cultivar fungus species – *Leucoagaricus gongylophorus* (Schultz & Brady, 2008). This is the most recently evolved cultivar fungus, even more recently than the ants that cultivate them (Mikheyev et al., 2010; Mueller et al., 1998). *L. gongylophorus* is distinguished from other cultivar fungi by its conspicuous and specialized morphological adaptations such as *gongylidia*, hyphal swellings that are the primary food source for the ants (de Fine Licht et al., 2014). Similarly, domesticated agriculture is conducted by the ant genera *Trachymyrmex* and *Sericomyrmex*, who cultivate a single clade of *Leucoagaricus* cultivar fungus and engage in markedly less cultivar sharing than basal and yeast agriculture (Mikheyev et al., 2010). The fungi cultivated in domesticated agriculture shares many of the morphological and evolutionary characteristics of leaf-cutter cultivar (de Fine Licht et al., 2014). Ants that conduct either leaf-cutter or domesticated agriculture are together known as “higher attines”.

The transition to domesticated and leaf-cutter agricultural types is marked by the higher attine fungal cultivar’s inability to live without their ant hosts and to form mushrooms, even under laboratory conditions (Weber, 1972). In this way, they never reproduce sexually and rely on their ant hosts to be transmitted to new colonies (Schultz & Brady, 2008). This transmission mode

suggests that these cultivar fungi naturally exist as asexual, clonal crops. One study detected only a single cultivar genotype within *Atta* fungus gardens sampled over several years (Mueller et al., 2010), implying that *Atta* maintains stable fungal monocultures within their fungus gardens. However, these experiments used microsatellite genotyping based on short tandem repeats (STR) - polymorphic loci composed of repeated sequences - to detect cultivar genotypes in a single fungus garden. This genotypic technique forces a single consensus from all template DNA sequences, and can therefore not detect heterogeneous multi-copy genes from a single sample. These cultivar samples were also collected directly from *Atta* colonies, which can extend to a diameter of 30 – 600 m^2 from the central mound and contain millions of ants housed in a complex network of multiple fungus gardens (Weber, 1966). The size of these colonies makes an exhaustive sampling of gardens an extremely daunting task, but one that is necessary to conclusively test for monoculture.

Additional complications may arise because closely related leaf-cutter ants exist in overlapping habitats (Weber, 1966). This makes horizontal exchange of symbionts during colony founding and symbiont theft after garden loss common ecological processes that influence cultivar diversity (Howe et al., 2018; Mikheyev, et al., 2007). This has been suggested to explain the lack of co-cladogenesis between cultivar fungi and their hosts (Poulsen et al., 2009). Crop diversity is also controlled within fungus gardens by the cultivar fungi themselves, such that domesticated fungi actively reject hyphal fragments from neighboring colonies (Poulsen & Boomsma, 2005). This antagonism, however, has not been tested across a wide array of cultivar genotypes. Additionally, these observations were made on cultured isolates grown in lab conditions and therefore may not represent natural conditions.

Leaf-cutter ants have excellent potential to explain the mechanisms that govern complex obligate mutualisms. However, the ecological complexities of this symbiosis can make it less suitable to study cultivar diversity and dynamics. Some domesticated agriculture species, such

as *Trachymyrmex septentrionalis*, may be more tractable for such studies. *T. septentrionalis* can be found throughout the Eastern U.S. and has comparatively small colonies that contain up to 1,000 ants (Rabeling & Johnson, 2007). This species only overlaps with other fungus-growing ant species at the extreme edges of its range, making host-cultivar swapping much less frequent (Mikheyev et al., 2008). This distribution covers a large climactic gradient throughout the Eastern U.S. Previous studies of these *T. septentrionalis* ants and their cultivar fungi described incongruences between the genetic structures of the cultivar fungi and their ant hosts (Mikheyev et al., 2008). Although *T. septentrionalis* ants split into eastern and western phylogroups, their cultivar fungi formed four distinct clades, Types A-D. Interestingly, the distribution of these clades did not correlate with geography. This diversity was attributed to biogeographical forces acting differently on the cultivar and their ant hosts, but does not explain how (Mikheyev et al., 2008).

These results indicate that clade-level diversity exists in a *T. septentrionalis* cultivar, despite its being thought to be clonal, vertically transmitted, and asexual. Additionally, these geographic data imply that these clades did not arise allopartically, and suggested some overlap of distinct cultivar types overlap between ant populations. However, these results were based on a multi-gene phylogeny generated using Sanger sequencing, which shares the same limitations as STR typing by forcing heterogeneous genetic data into a single consensus sequence. This method of pseudo-community analysis therefore does not resolve any genetic heterogeneity of cultivar fungi within each fungus garden, but rather generates a consensus of the predominant sequence types within each *T. septentrionalis* colony.

To accurately explain the evolution of genetic diversity within cultivar fungi, it is crucial to understand their sexual state and life cycle. For example, one recent genomic study of the fungal symbiont of the leaf-cutting ant *Acromyrmex echinatior* found strong evidence that more than two haploid genomes occurred within a single cultivar mycelium (Nygaard et al., 2016).

This observation was bolstered by a second study demonstrating that cultivar fungi raised by evolutionarily advanced ants (domesticated agriculture and leaf-cutting agriculture) were polykaryotic (Kooij et al., 2015). The degree of multinucleation found in cultivar fungi belonging to different agricultural levels co-varied with the colony size of their ant hosts (and therefore the degree of evolutionary development). These data suggested that cultivar symbionts of Leaf-cutter ants (like *Atta and Acromyrmex*) are highly polyploid; and domesticated agriculture ants (like species of *Trachymyrmex*, including *T. septentrionalis*) are lowly or facultatively polyploid. These data also suggest the cultivar could be heterogeneously haploid, though it was predicted that this model would be unstable without clamp connections. Together, these genomic developments are suggested to make the mutualism more productive and less vulnerable to infection from pathogenic fungi (Kooij et al., 2015).

Previous studies of cultivar genetics, evolutionary history, and patterns of diversity leave a gap in our knowledge of what genetic mechanisms generate cultivar diversity and how this influences the fungus-growing ant symbiosis. Capturing the total genetic diversity of the cultivar within fungus gardens is an important step towards accurately describing their evolution. Community amplicon sequencing coupled with clustering into amplicon sequence variants (ASVs) is particularly suited for such sensitive analyses (Callahan et al., 2017). Understanding how fungal symbionts influence the fungus-growing ant symbiosis depends on answering two questions about their evolutionary history. First, what is the genetic diversity of these fungi; is it greater than what we currently know? Second, what processes (i.e. genetic mechanisms) might explain these patterns of diversity?

The idea that *T. septentrionalis* rear a single, asexual cultivar fungus is widely accepted. Previous studies have also shown that *T. septentrionalis* cultivar ITS sequences form multiple clades, but the data supporting these ideas lacks the ability to characterize genetic diversity within individual fungus gardens. I therefore hypothesize that there is genetic diversity in *T.*

septentrionalis fungus gardens that has not yet been observed. Such diversity among these asexual fungi could be due to: i) polyculture, where *T. septentrionalis* fungus gardens contain more than one cultivar strain, or ii) polykaryotic or heterogeneously haploid cultivar fungi, where single cells contain multiple genetic variants in one or separate nuclei, respectively. All of these mechanisms would generate a diversity of ITS alleles. To test this hypothesis, we will create new ITS phylogenies using Sanger sequencing to characterize cultured isolates collected from throughout the Eastern U.S., to recapitulate the previous results and facilitate cultivar typing. We will then perform community amplicon sequencing, first on freshly-collected gardens to identify its total ITS allelic diversity, and also on cultured isolates to compare their genetic diversity to that of freshly-collected fungus gardens. We will observe the nuclear state of cultured isolates with fluorescent microscopy to determine if they are polykaryotic, and analyze the genetic stability of cultured isolates with community amplicon sequencing over multiple generations. Together, these experiments will reveal the genetic diversity within *T. septentrionalis* fungus gardens and suggest the genetic mechanisms by which it is maintained.

3 - Methods

3.1 - Sample collection

T. septentrionalis colonies were collected from sites throughout the Eastern U.S. (**Table 4.2.1**). Colonies were visually identified by their conspicuous crescent-shaped mounds and presence of *T. septentrionalis* workers, following previous descriptions (Rabeling & Johnson, 2007). A hole approximately 30-60 cm wide was dug 60cm from the concave aspect of the mound and excavated using a hand shovel. *T. septentrionalis* fungus garden chambers were usually found 10-25 cm below the surface but this varied greatly, approaching 70 cm in some colonies. Soil was carefully removed from the mound-facing surface until a chamber was revealed, which was carefully excavated to reveal the entirety of the chamber. Fungus gardens were retrieved using a sterilized spoon and placed into premade humidity-controlled boxes lined with plaster of Paris. Colonies were maintained in these boxes and fed sterilized cornmeal in the lab. Field-collected fungus garden samples were subsampled after excavation in DESS and immediately frozen on dry ice before being stored at -80°C in the lab (Lee et al., 2018).

Isolates of cultivar fungi were cultured from fungus gardens after colonies had adjusted to laboratory conditions and began incorporating cornmeal into their gardens (after approximately 2–4 weeks following collection). Fluffy hyphal extensions were taken from the fungus garden and grown on BD Difco potato dextrose agar. These were prepared using 39g of media per liter of sterilized water, to which 50mg/L of filter-sterilized penicillin and streptomycin were added after the media was autoclaved and cooled to ~50°C. Isolates were monitored and re-plated for purity until they were no longer visibly contaminated. Purified strains were grown for 4 weeks at 27°C.

For the cultivar subsampling and growth experiment, cultures were isolated from lab-reared gardens (labeled “generation 0”) as described above. These isolates were periodically

examined for visible contamination using a dissection microscope while they grew for two weeks at 27°C in a room temperature incubator. Visibly contaminant-free cultures were used to inoculate fresh media by removing a 5mm^2 agar core containing cultivar hyphae from the initial culture media. These first subcultures were labelled “generation 1.” After growing for two weeks, these subcultures were visually inspected for contamination and photographed beside a ruler. The total growth diameter for each subculture was measured from the extent of hyphal growth within the medium. Visibly contaminated subcultures were photographed and removed from the study. For all pure cultures, a 5mm^2 agar core was taken containing approximately half of the colony, used to inoculate fresh media, and labeled “generation 2.” The remaining fungal tissue from each generation 1 culture was used for DNA extraction. This process was repeated for 3 additional generations.

3.2 - Sample Processing

DNA was extracted from cultivar isolates using an adapted CTAB buffer protocol (Lee et al., 2018). Approximately 10mg of fungal tissue was taken from the top of each culture using sterile forceps. This protocol was modified to include only two 2-minute cycles of bead beating (Biospec Minibeadbeater) before DNA extraction. Each DNA pellet was washed with 1ml of 70% ethanol and centrifuged at 15,000xg for 1 minute to remove PCR-inhibiting molecules. This ethanol wash was discarded and the pellets were dried at 27°C for two hours before rehydration in nuclease-free water. Rehydrated DNA were quantified using a BioSpec Eon plate reader spectrophotometer with a Take3Trio plate. DNA extracts with A_{260}/A_{280} or A_{260}/A_{230} ratios of ≥1.5 were considered free of PCR-inhibiting molecules and used for PCR.

Partial ITS gene sequences from were amplified using ITS1 and ITS4 primers and DNA extracted from cultured isolate DNA (White et. al, 1990). Approximately 1.5µl of each DNA extract (<1,000 ng) was used as the template for a 25µl PCR reaction using New England Biolab’s Q5 High-fidelity PCR kit. Each PCR reaction contained 1X Q5 reaction buffer, 5X Q5

High GC Enhancer 5mM dNTPs, 15µM primers, 5mM $MgCl_2$, 0.375µl BSA, and 0.02U/µl Q5 High-Fidelity DNA Polymerase. PCR reactions were run on a T-100 Thermal Cycler (BioRad) with an initial denaturing step of 95°C for 3 minutes, 30 cycles of 30 seconds at 95°C, 45 seconds at 50°C, and 45 seconds at 72°C, followed by a final extension step at 72°C for 2 minutes. Bands were visualized using agarose gel electrophoresis to confirm the PCR product was approximately 800bp. All PCR products were cleaned using Agencourt AMPure XP (Beckman Coulter) magnetic beads following the manufacturer's protocol before submitting them to the University of Connecticut's Center for Genome Innovation (CGI) for Sanger sequencing.

The same protocol was used to extract genomic DNA for community amplicon sequencing. Genomic DNA was quantified and quality checked using the same method described above, but did not require cleaning with magnetic beads. Samples were diluted to contain <1,000ng of DNA in a 5µl solution and submitted to the University of Connecticut's Microbial Analysis, Resources, and Services (MARS) facility for community amplicon sequencing. PCR products were pooled for quantification and visualization using the QIAxcel DNA Fast Analysis instrument (Qiagen), normalized based on the concentration of DNA with sizes ranging from 350-420 bp, and then pooled using a QIAgility liquid handling robot. These pooled PCR products were cleaned using Mag-Bind RxnPure Plus (Omega Bio-tek) magnetic beads according to the manufacturer's protocol. The cleaned pooled products were sequenced on an Illumina MiSeq using a v2 2x250 base pair kit (Illumina, Inc). The protocol for DNA extraction, PCR screening, and community amplicon sequencing for wild garden subsamples collected in the field is previously described in Adams, 2017.

3.3 - Bioinformatic analysis

Sanger sequences of cultured isolates were processed with the staden package v2.00b6 (Staden et al, 2010) to assemble consensus sequences. These sequences along with related

reference sequences (Kooij et al., 2015; Mikheyev et al., 2008) were then aligned with MUSCLE v3.5 (Edgar et al, 2004) to create a multiple sequence alignment. These alignments were trimmed to remove primers but were not processed to not mask polymorphic sites. The best DNA substitution models for these ITS sequences were estimated using jModelTest v2.1.10 (Posada, 2008), and the best substitution model for our data was GTR + I + G. All phylogenies used this substitution model and were computed using RaxML v8.2.4 (Stamatakis, 2014) with a rapid bootstrap analysis for 500 inferences before searching for the best-scoring minimum-likelihood tree.

Community amplicon sequences were demultiplexed using mothur 1.35.1 (Schloss et al., 2009). ITS gene sequences were processed in R v3.4.3 (R Development Core Team, 2017) using the DADA2 v1.7.0 (Callahan et al., 2016) pipeline (https://benjjneb.github.io/dada2/tutorial.html). Processed sequence and metadata files were imported into phyloseq v1.22.3 (Mcmurdie & Holmes, 2013) and checked for the occurrence of contaminants using decontam v0.20.0 (Davis et al, 2018). Negative controls for freshly-collected fungus garden samples returned no sequences, so decontam was not used to process these data. For cultured isolates, contaminants were screened using the “prevelance” protocol with the P* threshold set at 0.5. Because all sequences in the negative controls were not significantly prevalent in the cultured isolate samples, no sequences were removed. Once taxa were matched to sequences within the UNITE database (Nilsson et al., 2019), all samples that were not classified to family-level within *Agaricacae* were considered non-cultivar fungi. The resulting phyloseq objects were subsampled to 1000 reads and their Amplicon Sequence Variant (ASV) counts were converted into relative abundances. Phyloseq was also used to calculate alpha diversity (Shannon & Chao1) using the estimate_richness function. Processed data files were then exported into R for statistical analysis. The hierarchical clustering analysis heatmap of the pairwise similarity between ITS ASVs in freshly-collected gardens was made using BLAST

v2.2.28 (Camacho et al., 2009) to compare all ASVs to each other and to the ITS sequences used to create the cultured isolate ITS phylogeny, from which cultivar Types were defined. Co-occurrence of ITS ASVs in freshly-collected gardens was calculated using SparCC (Weiss et al., 2016). The aov command was used to perform analysis of variance (ANOVA) testing. The code to reproduce these analyses is available in **File S1.**

3.4 - Microscopy

I observed the morphology of cultivar isolate cells using a compound light microscope (Amscope). Small tufts (less than 1mm^2 diameter) were taken from culture plates and wet-mounted on a slide with 0.1% (w/v) aqueous congo red solution (3,3′-([1,1′-biphenyl]-4,4′-diyl)bis(4-aminonaphthalene-1-sulfonic acid). For nuclei staining, small cultivar tufts were first added to an 80% (v/v) aqueous glycerol solution and incubated at 0°C for 30 minutes. These solutions were mixed using a 20µl micropipette and vortexed periodically to agitate and separate hyphal masses. A 10µg/ml solution of 4′,6-diamidino-2-phenylindole (DAPI) was applied to wet-mounted fungal tufts, covered with a cover slip, and incubated for 30 minutes in darkness (Villa et al., 2009). After incubation, these slides were washed with a 1X phosphate buffer solution (PBS: 8g NaCl, 0.2g KCl, 1.44g Na_2HPO_4, and 0.24g KH_2PO_4, pH 7.4) to remove remaining dye. Cover slip edges were sealed with clear acrylic nail polish to prevent cell drift during microscopy. Prepared slides were viewed on a Zeiss Axiovert 200M inverted fluorescence DIC microscope. Images of hyphae stained with DAPI were captured in duplicate under bright field and a fluorescent filter for DAPI at 63x, and captured across a Z-plane at approximately every 0.25µm and stacked. Images were processed using ImageJ 1.52a (Schneider et al., 2012) to synchronize the bright field and fluorescence images and increase contrast to clearly distinguish nuclei and reduce noise from autofluorescence. Using cell wall boundaries visible in the bright field images, nuclei were manually counted to quantify the number of nuclei per cell. Nucleus size was quantified in ImageJ based with the diameter of

each nucleus measured by the pixels it occupied in the image. This was then converted into a nucleus diameter in micrometers.

4 - Results

4.1 - Culture-based phylogeny

Previous studies have established that *T. septentrionalis* ants rear cultivar fungi related to those cultivated by other members of the Attine domesticated agriculture group. To confirm the phylogenic placement of the *T. septentrionalis* cultivar isolate sequences generated by our study, we created a new ITS phylogeny of our cultivar isolates and those previously described from other ant agriculture types (Kooij et al., 2015). This phylogeny showed that each agricultural type grouped together, consistent with previous studies (**Figure 4.1.1**). The coral, basal, and yeast agriculture groups all clustered closest to the outgroup, representative of their cultivation by attini belonging to lower agricultural groups. The domesticated agriculture cultivar clade formed a sister clade to fungi from the leaf-cutting group. Sequences from our *T. septentrionalis* cultivar isolates were nested within the domesticated agriculture group in several distinct clades. Collectively, these data indicate the cultured isolates from this study were phylogenetically related to other previously-described *T. septentrionalis* cultivar isolates within the domesticated agriculture group. These sequences do not form a monophyletic clade within this group, consistent with a previous study (Mikheyev et al., 2008).

To more closely investigate the phylogenetic placement of our isolate sequences, we recreated the phylogeny of (Mikheyev et al., 2008) (**Figure 4.1.2**) to use as the foundation for a new phylogeny that also includes our new *T. septentrionalis* cultivar isolate ITS sequences (**Figure 4.1.3**). Isolate ITS sequences from our study were classified within all of the groups described by Mikheyev et al. 2008 except for their cultivar Type D, which was represented by a single sample (TV011102-02) in their study. The addition of our sequences broadened the existing polytomy within Type A and Type B cultivar clades, but created fairly resolved groups in

the Type C clade. Another well-resolved group of cultivar sequences from our study formed a sister clade to the Type A and Leaf-cutter cultivar clades. The samples within this group (JKS002428, JKS0002427, and JKS0002332) did not group with any sequences from the previous study, and so we label this clade here as “Type E”. We therefore conclude that our cultivar isolate sequences can be categorized within clades previously established, with the exception of the novel Type E clade.

4.2 - Community amplicon sequencing of freshly-collected fungus gardens

The culture-based phylogeny above confirmed there is clade-level phylogenetic diversity among the cultivar fungi isolated from *T. septentrionalis* fungus gardens collected throughout the Northeastern U.S. However, we found that different cultivar types were present in the same geographic location, consistent with (Mikheyev et al., 2008) (**Figure 4.1.3**). To better understand this diversity and distribution of *T. septentrionalis* cultivar fungi, we used community amplicon sequencing to characterize all ITS gene sequences present within each fungus garden. The resulting ITS amplicon sequence variants (ASVs) indicated that freshly-collected gardens are predominantly comprised of the cultivar fungus, but also contain a variety of non-cultivar fungi (**Figure 4.2.1**). The reads within this category originate from a broad diversity of fungi, most of which are saprotrophic species that naturally occur in soils (Adams, 2017). A small subset of samples contained more non-cultivar sequences than cultivar sequences, with a maximum of 88% read abundance. Most gardens were composed of ≤2% non-cultivar reads, as anticipated. All freshly-collected gardens, including those dominated by non-cultivar reads, were incorporated into subsequent datasets to investigate cultivar ITS ASV diversity in freshly-collected gardens.

After removing non-cultivar reads, we found that *T. septentrionalis* fungus gardens contained a diverse mixture of cultivar ITS sequence variants of various abundances (**Figure 4.2.2**). The ASVs present in each *T. septentrionalis* fungus garden varied from state to state and

slightly between sites within the same state. Collectively, these data suggest that freshly-collected gardens are polycultures, containing the alleles of what appear to be multiple strains of cultivar fungi. There is a much higher abundance of ITS ASV alleles is present in gardens than previously acknowledged, which might dramatically impact the structure of cultivar isolate ITS gene phylogenies (such as that in **Figure 4.1.2**).

Our community amplicon sequencing and single gene phylogenies produced very different representations of the fungal communities *within T. septentrionalis* fungus gardens. We therefore attempted to classify the community amplicon ASVs into the cultivar types determined from **Figure 4.1.3** to reconcile these views of community structure. However, a phylogenetic analysis of these short ITS sequences had very poor resolution and too many unresolved clades to reliably classify cultivar ASVs (data not shown). Instead, we compared the percent identity of all freshly-collected garden ITS ASVs and the reference sequences from **Figure 4.1.3** with a basic local alignment (BLAST) and clustering analysis. The resulting heat map clustered ASVs according to their percent identity to each other (**Figure 4.2.3**), and allowed us to assign all ITS ASVs to *T. septentrionalis* cultivar types based on their similarities to the reference sequences (**Figure 4.1.3**). Cultivar ASVs that clustered with Type A reference sequences formed a well-resolved group that was consistent with **Figure 4.1.3**. Many other cultivar ASVs grouped with Type B cultivar reference sequences, albeit with greater heterogeneity compared to the Type A group. This may indicate that the Type B cultivar clade is a more variable “grade” rather than a true clade. Cultivar ASVs also grouped with Type C and Type E cultivar reference sequences in lower numbers than the for Types A and B, but none grouped with the Type D reference sequence from Mikheyev et al. 2008. Some ASVs clustered with the outgroup reference sequences and were removed from subsequent analyses, but none grouped with the Leaf-cutter cultivar reference sequences. A small proportion of cultivar ASVs did not cluster clearly with any reference sequence type but were still clearly cultivar ITS sequences; these

were labelled as Unclassified. Collectively, these data classified nearly all cultivar ASVs within the groups established by the ITS gene phylogenies generated using cultured *T. septentrionalis* cultivar isolates.

With the cultivar ASVs from freshly-collected gardens classified into types, we wanted to investigate how cultivar types were distributed geographically. Additionally, we questioned whether multiple types could be found within single gardens, just as multiple ASVs had been found in **Figure 4.2.1**. We found that Type B cultivar sequences were the most widely distributed and highly abundant type, constituting more than 95% of the total ITS reads in most of the fungus gardens in all states sampled (**Figure 4.2.4**). Type A cultivar sequence abundance was more variable, making up 2%-65% of the reads in freshly-collected gardens from New Jersey, North Carolina, Georgia, and Florida. Unclassified cultivar sequences were found in low abundances (<%2) in fungus gardens collected in New Jersey, North Carolina, and Georgia, and Type C sequences were found in similar abundances (<2%) in samples from New Jersey, North Carolina, and Florida. Type E sequences were only found in extremely low abundance (<1%) in freshly-collected fungus gardens collected in North Carolina, and Type D cultivar sequences were not found in any of the freshly-collected gardens that we sampled. Collectively, these data suggest that Type B cultivar sequences are the most dominant and widely distributed throughout the Eastern USA, and that Type A sequences also occur in moderate-to-high abundance in all states except for New York and Louisiana. Freshly-collected gardens also can contained sequences from multiple cultivar types, further suggesting that freshly-collected gardens are polycultures.

4.3 - Community amplicon sequencing of cultured isolates

The idea that fungus gardens are a polyculture of cultivar fungi directly contradicts previous studies that show the stabile colonization of a single cultivar strain over years of sampling from the same garden (Mueller et al., 2010). There exists, however, the possibility that

the polyculture that we observed could be caused by multiple allelic variants occurring within single cells, as opposed to originating from two separate cultivar strains coexisting within the same fungus garden. Such cells have been inferred by previous studies indicating that polykaryotic or heterogeneously haploid cultivars are grown by highly evolved fungus growing ants (Kooij et al., 2015). To rule out cultivar strain polycultures as a potential mechanism for our observed genetic diversity in *T. septentrionalis* fungus gardens, we isolated cultures of cultivar from lab-adapted gardens and used community amplicon sequencing to identify the ASVs present within them. We hypothesized that if freshly-collected gardens were polycultures of multiple cultivar strains, we would only observe the ITS ASVs of a single, or very few, alleles within cultured isolates.

We first surveyed the ratio of cultivar to non-cultivar reads in our cultured isolates, which did not differ greatly from that of freshly-collected garden samples (**Figure 4.3.1**) The reads within this category consisted primarily of the same classes of fungi found in freshly-collected gardens as described above. Most isolates contained ≤2% non-cultivar fungi, while one isolate from a Georgia colony contained ~6%, and one isolate from North Carolina contained ~95% non-cultivar reads.

After removing non-cultivar sequences, each cultivar isolate still contained several ITS ASVs. These differed slightly from that of the freshly-collected gardens from the same geographic origin (**Figure 4.3.2**). Most cultivar ASVs in our cultures belonged to cultivar Type B, similar to freshly-collected gardens (**Figure 4.3.3**). The dominant ASVs in our cultures were similar to those in corresponding freshly-collected gardens, but some cultures (such as WBU, which was dominated by cultivar Type E reads) were dominated by ASVs that had low abundances in all freshly-collected gardens. Some cultivar cultures, such as LRS and AWF, contained ASVs from up to three cultivar Types. Freshly-collected fungus gardens had higher alpha-diversity of ASVs compared to cultured isolates (**Figure 4.3.4**). This comparison,

however, included all freshly-collected gardens and cultured isolates, which does not allow for direct comparisons. We therefore correlated the alpha diversity of freshly-collected gardens to isolates cultured from those gardens to better quantify if alpha diversity changes following culture. We hypothesized that if there was no change in alpha diversity between these conditions, this plot would have an equation of y = x, or a slope of 1. In comparison, freshly-collected gardens had a higher alpha diversity than the isolates cultured from them (y=0.3581x + 5.9774) (**Figure 4.3.5**).

Collectively, these data suggest two things. First, cultured isolates still contain a mixed, albeit reduced diversity of ASVs. We therefore conclude that freshly-collected gardens are not polycultures of multiple cultivar strains, but that the cultivar itself can contain sequences of multiple ITS ASVs, even from different cultivar types. This more closely matches previous hypotheses of a polykaryotic/heterogeneously haploid cultivar, but cannot adequately determine which of these possibilities is more likely. Second, culturing isolates from freshly-collected gardens appears to enrich for some ASVs over others. Because much of previous work on cultivar diversity is based on ITS gene phylogenies of cultured isolates that were characterized by Sanger sequencing, we suggest that these may have experienced enrichment bias that has altered the structure of these phylogenies in unknown ways. We believe this is reasonable to conclude based on the limited occurrence of Type E cultivar sequences in most freshly-collected garden samples, which became enriched to entirely Type E reads in JKA004259. This would also explain why Type D reads were never recovered in any freshly-collected garden samples or cultured isolates, and why this sequence type is represented by a single isolate sample in the previous study.

4.4 – Co-occurrence of ASVs

Though the cultured isolates contain a mixture of different ASVs and cultivar types, measurements of alpha diversity alone cannot untangle the genetic mechanisms that created

this diversity. One way to examine this, however, is by considering the intensity of genetic linkage that would exist in each of these mechanisms. We therefore tested the correlations between the abundance of all ASVs in all freshly-collected gardens samples. We hypothesized that if the cultivar was homogenously polykaryotic (or if differently multi-copy ITS genes occurred in the same genome), many genes would have a high degree of correlation because they occur in the same *non-recombining* genome. We found that no cultivar ASVs co-varied negatively in this dataset, indicating a lack of competition between cultivar ASVs within *T. septentrionalis* fungus gardens (**Figure 4.4.1**). Some clusters of ASVs weakly correlated, but none of these correlations strongly indicated direct genomic linkage of the alleles. We therefore conclude that the cultivar fungus is not likely to be homogeneously polykaryotic based on the lack of strong allelic linkage, but that it may potentially be heterogeneously haploid or a mixed polyploid.

4.5 – Microscopy of cultured isolates

Community amplicon sequencing showed that cultured *T. septentrionalis* cultivar isolates contained a diverse ITS sequence variants, sometimes including sequences from multiple cultivar types. These genetic differences might be lead to phenotypic differences between cultivar strains. We therefore investigated if isolate strains differed in histology. Cell dimensions, shapes, and connections to other cells varied considerably between cultured isolates (**Figure 4.5.1**). Morphologically-distinct cell types occur both within the same cultivar culture (e.g., vegetative cells, staphylae/gongylidia; clavulate gongylidia, irregular hyphae) (**Figure 4.5.1**). The density of gongylidia within staphylae and the whether gongylidia only occur within staphylae also varies between isolates. Phenotypic differences extend to macroscopic characters such as where staphylae form, how densely packed they are with gongylidia, and macroscopic colony morphology (**Figure 4.5.2**). Collectively, these data suggest there is a high degree of morphological diversity within cultured isolates, ranging from vegetative cells to the

highly evolved, specialized cells and organs for feeding the *T. septentrionalis* hosts. We also observed that clamp connections – a cellular feature that fungi use to equally distribute heterogeneous nuclei between cells after mating – occur abundantly between hyphae in our cultivar isolates. These clamp connections provide evidence that the cultivar fungus actively distributes nuclei between cells as their hyphae grow.

The morphological diversity of isolates and evidence of nuclei-shuffling structures suggests that different cells in a single fungus garden might contain different genomic compositions. If true, this would explain why we observed diverse ASVs in our cultured cultivar isolates. Previous studies have suggested that fungi grown by evolutionarily advanced hosts such as *T. septentrionalis* are polyploid and multinucleated (Kooij et al., 2015). To test if our *T. septentrionalis* cultivar isolates match this polykaryotic hypothesis, we used fluorescent microscopy to survey the number of nuclei present in each cell. Cells from our cultivar cultures were highly multinucleated **Figure 4.5.3)** and contained nuclei of various sizes ranging from 0.25µm – 2.5 µm (**Figure 4.5.4**). The average number of nuclei per cell in lab-adapted colonies JKH000062, JKH000130, and JKH000144 were 10 ± 4, 6 ± 5, and 15 ± 13 nuclei per cell, respectively (**Figure 4.5.5**). In all strains, micronuclei (with diameters ranging from 0.4 µm – 0.6 µm) occurred in the highest abundance, while nuclei resembling standard sizes (around 1 µm) (Kooij et al., 2015) were the second most abundant (**Figure 4.5.4**). These data corroborate the nuclei counts for the domesticated agriculture hosts *T. septentrionalis* and *Sericomyrmex amabilis* in (Kooij et al., 2015)that reported an average of 10.80 ± 0.43 nuclei per cell. We therefore conclude that our cultured isolates are polykaryotic or heterogeneously haploid.

4.6 – Genetic stability of cultured isolates across generations

If the *T. septentrionalis* cultivar is polykaryotic, we would expect its genomic content to be more stable than if it is heterogeneously haploid, because all alleles would occur on the same genome of a single nucleus rather than across separate haploid nuclei. We therefore

designed an experiment to test the genetic stability of multigenerational lineages of cultured cultivar isolates, tracking growth rates, mortality, and genetic composition via ITS gene community amplicon sequencing for each generation. In total, 17 cultivar isolate lineages were cultivated from three lab-adapted fungus gardens: JKH000062, JKH000130, and JKH000144, with 8, 4, and 5 lineages, respectively. A genealogical tree of these lineages shows the experimental setup and total growth area for each lineage (**Figure 4.6.1**). We first tracked the growth rate and mortality of each multigenerational lineage to assess the health and productivity of the isolates through time. We hypothesized that under a polykaryotic model, the growth rate of isolates and mortality would not change significantly over time because they would not lose allelic diversity. The growth rate of our cultivar isolate lineages differed significantly between generations (ANOVA, $F(3,32) = 6.04$, $p = 0.002$) (**Figure 4.6.2**). A pair-wise posthoc comparison identified the growth rate of generation 1 differed significantly from that of generations 2 and 3 ($p \leq 0.005$). This suggests that the isolates experienced a significant reduction in growth rate after generation 1, which was the most productive generation. The most significant mortality occurred in generations 1 and 3 (with mortalities of 41% and 33%, respectively). This suggests that many isolates in generations 1 and 3 may have had unstable combinations of allelic variants and experienced mortality as a result. Collectively, these data suggest that the productivity and stability of the isolates decreased significantly after generation 1, and experienced instability and reduced productivity thereafter.

Although these data show that growth of the isolate lineages changed between generations, growth rate is only a correlate of genetic stability. To assess genetic stability, we used community amplicon sequencing to track the ITS gene ASV composition of our cultivar isolate lineages across generations. We hypothesized that under a polykaryotic model, the ASV composition of each cultivar isolate lineage would remain constant between generations of sub-culturing because each ASV would be have direct genetic linkage to each other. If the isolates

were heterokaryotic, however, we would anticipate that ASV compositions might vary because of reduced linkage between them. We found that the ASV compositions of each cultivar isolate lineage: i) had lower alpha diversity compared to the lab-adapted gardens they were cultured from and ii) varied from the composition of the source fungus gardens and also slightly between generations of the same lineage (**Figure 4.6.3**). The ASV composition of our cultivar isolate lineages occupied three different, semi-stable states, represented by isolates 144_1_1, 160_1_1, and 219_1_1 (**Figure 4.6.3**). Interestingly, the ASV composition of some isolate lineages resembled that of isolates from other host gardens (e.g. 144_1_8 resembles JKH000219 lineages). These data corroborate our previous findings of the reduced alpha diversity and enrichment bias that occurred in cultured cultivar isolates compared to freshly-collected fungus gardens. This was especially apparent in the lineages from JKH000160, where six dominant ASVs occurred in the lab-adapted garden but only three occurred in later generations of its corresponding cultivar isolate lineages. Second, the composition of ASVs in isolate lineages was somewhat, but not completely, flexible. These data suggest that cultivar isolates are not polyploid, unless some degree of heterogeneity occurs. We therefore conclude this data more likely describes a heterogeneously haploid model, in which some degree of genetic linkage exists.

5 – Discussion

5.1 – New strategies to accurately assess fungal diversity in *T. septentrionalis* fungus gardens and cultured isolates.

All previous cultivar phylogenies have been constructed using single or multi-gene Sanger sequences. For example, (Mikheyev et al., 2008) created their *T. septentrionalis* cultivar phylogeny using multi-gene sequences from fungus gardens that were genotyped using microsatellite short tandem repeats (STR). Their results indicated that *T. septentrionalis* fungus gardens hosted different cultivar types that did not correlate with their geographic origin. Sanger sequencing creates strong consensus sequences with long read lengths from homogeneous samples, and has been widely applied for phylogenetic analysis of fungi, especially by sequencing the internal transcribed spacer region (ITS) gene (Schoch et al., 2012). However, Sanger sequencing condenses any genetic diversity in these samples into a single consensus sequence. This makes it incapable of detecting heterogeneous alleles within a single sample, which it instead presents as a single sequence with polymorphic sites. Polymorphic sites are typically considered during multiple sequence alignment, and it is common practice to mask these sites to avoid errors in the resulting phylogenies (Di Franco et al., 2019; Singh & Bhatia, 2016). This reduces the number of informative sites available to differentiate these sequences. *T. septentrionalis* ITS genes can differ by only a few bases, reducing the phylogenetic resolution of these groups by relying on a few substitutions to distinguish lineages. This can produce a phylogeny with unresolved polytomies and weak bootstrap values between branches.

To resolve these polymorphisms, we used community amplicon sequencing followed by clustering into ASVs to exactly resolve our fungus garden ITS gene sequences down to single nucleotide differences (Callahan et al., 2017). This method treats all polymorphic sites as distinct sequence variants, generating a high-resolution profile of sequence diversity within our

samples. The short sequences produced using this technique have lower phylogenetic resolution compared to Sanger sequencing, as evidenced by the convoluted phylogeny that we produced using these sequences (data not shown). We therefore applied community amplicon sequencing to profile total ITS gene allele diversity in freshly-collected *T. septentrionalis* gardens but not to resolve the phylogenetic relationship between cultivar alleles.

Both freshly-collected gardens and cultured isolates contained diverse and varied ITS alleles (**Figures 4.2.2 and 4.3.2**), indicating allelic diversity in both *T. septentrionalis* fungus gardens and cultivar isolates that has not yet been reported. The diversity of ASVs in cultured isolates highlights two important considerations. First, conventional techniques for isolating cultivar strains from lab-adapted fungus gardens as described in previous studies (Poulsen & Boomsma, 2005) (Khadempour, personal correspondence, 2018) do not result in mono-allelic cultures, but rather subcultures of a genetically diverse fungal community. This is evidenced by the mixed representation of ITS ASVs in our cultured isolates, and by the abundance of non-cultivar reads present in our isolates. The second, and perhaps most problematic, consideration from our data is that culturing can introduce enrichment bias compared to the fungus gardens that these cultures represent. For example, cultured isolate WBU contained mostly the novel Type E cultivar sequence, which were found in low abundances in freshly-collected gardens (**Figures 4.2.2 and 4.3.2**). This isolate was therefore enriched in sequences that were rare in their source fungus garden. We suspect this may also be true for the strains in Mikheyev et al's Type D clade (Mikheyev et al., 2008) which was represented by only one isolate in their study that matched none of our ASVs from this study. The influence that such enrichment bias on previous studies is unknown, but may have impacted the inferences that were drawn from these analyses. For example, previous studies have reported four widespread *T. septentrionalis* cultivar types, similar to our culture-based results, whereas our ASV analyses indicated that only one or two dominated throughout much of the Eastern U.S (**Figure 4.2.4**). Incorporating

community amplicon sequencing data therefore can reveal genetic diversity that was not captured by relying on culture-based phylogenies alone.

In light of this, what methods should we use to accurately assess and interpret the sequence diversity in *T. septentrionalis* fungus gardens and cultured isolates? We suggest two main priorities: i) overcoming sampling bias to capture the complete allelic diversity within the samples; and ii) classifying these data accurately within established phylogenies. The first objective can be achieved using community amplicon sequencing of freshly-collected gardens and clustering of those data into ASVs, as described above. Calculating a maximum-likelihood tree using these ASVs, however, does not resolve the sequences into cultivar types due to their short length. For this reason, we used an alignment-based classification using the cultivar Sanger sequences as representatives of each cultivar type (**Figure 4.2.3**). This differs from phylogenetic analysis in that it classifies sequences based on their highest scoring alignment to a query sequence and does not attempt to establish lineages or reconstruct an ancestral state. This allows for a higher-resolution clustering that is based on the differences between a few informative sites. This demonstrated that most of our ASVs could be clustered into cultivar types, excepts for a few that did not group with any cultivar type or outgroup sequences and were considered to be unclassified.

Collectively, our hybrid approach that first used long Sanger sequences to reconstruct cultivar types and then used these types to classify the shorter ASV data allowed us to accurately characterize the phylogenetic diversity of cultivar ITS gene sequences in *T. septentrionalis* fungus gardens and cultivar isolates. These data can be used to infer processes that maintain genetic diversity in the *T. septentrionalis* cultivar, which would have been impossible using traditional methods based solely on Sanger sequencing. This robust method of analysis should be more commonly applied in future studies of fungal diversity.

5.2 – Are *T. septentrionalis* fungus gardens a monoculture?

Our analyses were designed to answer the question: is fungal diversity in *T. septentrionalis* fungus gardens greater than what previous studies have demonstrated? The assumption that *T. septentrionalis* ant rear an asexual fungal symbiont is based on extensive testing which revealed that cultivar fungi of high attini cannot produce fruiting bodies (Weber, 1972). The assumption that fungus gardens are monocultures is largely borrowed from studies of highly evolved fungus-growing ants such as *Atta* (Mueller et al., 2010; Poulsen & Boomsma, 2005) . However, phylogeographic analyses of *T. septentrionalis* populations across the Eastern U.S. suggest that there are many different cultivar monocultures that coexist within larger populations of their hosts (Mikheyev et al, 2008). This diversity might imply that the composition of fungus gardens or cultures across larger populations can significantly influence their function. We hypothesized that genetic or ecological mechanisms (i.e. polyculture gardens, or a polyploid/heterogeneously haploid cultivar) may help explain the patterns of diversity observed in previous studies. The patterns of allelic diversity we observed matched what we would expect for i) a polyploid cultivar where alleles are not homogeneously distributed throughout the fungus garden – also known as a “low polyploid” (Kooij et al, 2015); or ii) a heterogeneously haploid cultivar with a low amount of allelic linkage. This is consistent with our observation of clamp connections between *T. septentrionalis* cultivar cells, which might transfer distinct nuclei between cells and lead to heterokaryosis. This possibility was downplayed by (Kooij et al., 2015), who did not observe clamp connections in these fungi.

We first collected Sanger ITS sequences from cultured isolates to produce a new *T. septentrionalis* cultivar phylogeny using isolates collected from across the Eastern U.S (**Figure 4.1.3**). This phylogeny acted as an initial survey of *T. septentrionalis* cultivar diversity across this geographic range, recapitulated the results of previous phylogenies, recovered a novel cultivar

type (Type E), and generated a reference phylogeny for the classification of our community amplicon sequence ASVs. Surprisingly, community amplicon sequencing of freshly-collected *T. septentrionalis* fungus garden samples revealed a high diversity of ITS ASVs within all freshly-collected fungus gardens (**Figure 4.2.2)**, suggesting that each *T. septentrionalis* fungus garden contained multiple ITS alleles. These were sometimes classified into multiple cultivar types within the same garden (**Figure 4.2.4)**. These data also prove that fungus gardens of *T. septentrionalis* are far more diverse than suggested by previous studies. More importantly, these data Sanger sequencing does not accurately represent diversity within fungus gardens. For instance, the diversity of cultivar types as described by (Mikheyev et al., 2008) suggests that fungus gardens across populations of *T. septentrionalis* can contain up to four cultivar types. In contrast, our ASV data reports that a single cultivar type dominates most fungus gardens across our collection range, with the exception of a few that represent multiple cultivar types (**Figure 4.2.4**). Our results, therefore, would more closely resemble a monoculture of cultivar types across a broad geographic range than previous studies. When ASVs are not classified to type, the data suggest fungus gardens are extremely diverse (**Figure 4.2.2**), representing variants of many cultivar alleles and appearing as a polyculture.

We also used community amplicon sequencing to characterize cultured *T. septentrionalis* cultivar isolates from lab-adapted fungus gardens. We hypothesized that if the fungus gardens were polycultures, each of these isolates would contain only a single ITS gene ASV. Our data shows that isolates contained many ASVs (**Figure 4.3.2**), albeit fewer than in freshly-collected *T. septentrionalis* fungus gardens (**Figures 4.3.4 and 4.3.5**). This suggests that fungus garden polyculture cannot explain the allelic diversity we observe, but that these allelic variants occur on a cellular level in the cultivar. Our SparCC analysis (**Figure 4.4.1**) demonstrated only weak correlations between ASVs in freshly-collected *T. septentrionalis*

fungus gardens suggesting that these ASV are not genetically linked, as would be expected if they occurred as allelic variants on the same genome without recombination.

We then tested the genetic stability of out cultivar isolates over multiple generations. Our results showed that ASV abundances were maintained between cultivar generations, although the diversity of these cultured strains was both less than in their source fungus gardens and changed between different ASV profiles in some lineages (**Figure 4.6.3**). These observations suggest that not all allelic variants are always needed to establish a stable genetic state for the cultivar. It also confirms, however, that some basal level of allelic variation is required for a stable genetic state. The composition of allelic variants in isolate lineages varied between the origin of each isolate, but these always established stable states of certain allelic variants that only varied weakly between generations. This could suggest that stable states represent compatible sets of allelic variants, implying that not all combinations of alleles are compatible. This is consistent with previous studies which report antagonistic interactions between genetically distinct symbionts in fungus gardens (Poulsen & Boomsma, 2005), and suggests that some degree of epistasis between incompatible alleles could explain the limited number of stable states that we observed.

These ASV data strongly suggest that different allele combinations may be present in different nuclei, consistent with (Kooij et al., 2015) . We therefore counted the number of nuclei in each cultivar isolate cell using fluorescent microscopy, and found that our isolates were highly polykaryotic and contained nuclei of various sizes (**Figures 4.5.3, 4.5.4, and 4.5.5).** Our microscopy results also suggested possible mechanisms by which cell-to-cell heterogeneity is maintained in the *T. septentrionalis* cultivar (**Figure 4.5.1**). These include hyphal anasmotoses (sites of hyphal fusion where nuclei from separate cells mix in a continuous cytoplasm) and clamp connections (specialized hyphal connections that can reshuffle heterogeneous nuclei between growing hyphae). The high abundance of these cell structures in our cultured isolates

implies that they engage in processes which deliberately lead to cell-to-cell heterogeneity by placing multiple heterogeneous nuclei within a continuous cytoplasm. We therefore conclude that these processes are consistent with the heterogeneous patterns of diversity we observed in our community amplicon sequencing.

Our experiments show mixed indications regarding which of genetic mechanism most accurately describes the allelic diversity in freshly-collected *T. septentrionalis* fungus gardens and cultured isolates. A polyculture garden is supported by the diversity of ASVs in wild gardens (**Figure 4.2.2**), but is refuted by high ASV diversity being maintained in cultured isolates (**Figure 4.3.2**), unless polyculture is interpreted at a cellular level. We confirmed that the cultivar is polykaryotic by florescent microscopy (**Figure 4.5.3**), a basic condition for polyploid and heterogeneously haploid states. Polyploidy would explain the stability of ASVs observed in cultured isolates across generations (**Figure 4.6.3**). However, if all allelic variants were contained within a single nucleus, they would share a high degree of linkage (**Figure 4.4.1**). Polyploidy also cannot explain the reduction of diversity between freshly-collected fungus gardens and cultured isolates (**Figure 4.3.5**) unless we assume that there are heterogeneous polyploid nuclei, or predict alternative stable compositions of ASVs unless there is epistasis that only allows for certain ASV combinations. However, neither of these assumptions explain how cultured isolates enrich allele variants of cultivar types (such as Type E; **Figure 4.3.2**) that are comparatively rare in freshly-collected gardens.

A heterogeneously haploid cultivar can explain why genetic diversity decreases between lab-adapted fungus gardens and cultured isolates: heterogeneous alleles are not homogeneously dispersed, leading to the recovery of a subsample of total garden diversity in cultures. This would imply, though, that subsampling across generations would lead to a decrease in alpha diversity throughout generations, which is not something that we see (**Figure 4.6.3**). However, it is still compatible with the results of our correlation analysis (**Figure 4.4.1**) in

that the alleles are not strongly correlated with one another because they occur in separate nuclei. This would also accommodate the wide diversity of allelic composition we see in freshly-collected gardens (**Figure 4.2.2**), and would corroborate the hypothesis of (Kooij et al., 2015). However, it provides little explanation for the occurrence of the stable ASV compositions we observed in the genetic stability experiment (**Figure 4.6.3**), unless some amount of epistasis occurs between the alleles to promote certain allelic combinations.

Direct linkage between allelic variants in heterogeneous nuclei would explain the different stable ASV compositions in cultured isolates across generations. It would also explain why cultured isolates always retain diverse ASVs. However, the co-occurrence of ASVs did not imply that any alleles are strongly correlated (**Figure 4.4.1**). This conclusion is further complicated by the fact that stable ASV compositions were flexible and cannot explain how rare allelic variants can become enriched. However, if we assume that conditions for recombination between linked allelic variants on heterogeneously haploid nuclei are evident in our microscopy observations, the low degree of correlation and flexibility of stable combinations of ASVs can be considered as evidence of recombination between these nuclei.

We therefore conclude that the cultivar displays genetic patterns best resembling heterogeneous haploidy, in which genetic homogeneity is maintained through some degree of genetic linkage but not absolutely maintained. This can be explained by the occurrence of clamp connections and anastomoses in cultivar isolates, which might shuffle nuclei within a continuous cytoplasm and present opportunities for recombination to occur.

5.3 - Significance and implications for fungus-growing ant symbioses.

Our results demonstrate that *T. septentrionalis* cultivar fungi maintain a high degree of allelic diversity, even though they are not known to sexually reproduce via conventional pathways (Weber, 1972). However, these data do suggest that some recombination or shuffling of nuclei occurs between cultivar cells. Normally, a high diversity of ASVs such as what we observed would be described as a community, although fungus gardens have been considered monocultures (Mueller et al., 2010). In light of these contradictions, we consider *T. septentrionalis* cultivar fungi to be “genetic populations, although morphologically single entities”, as others have described certain arbuscular mycorrhizal fungi (Krishna, 2005).

The high allelic diversity that we observed in our community amplicon sequencing data of freshly-collected *T. septentrionalis* fungus gardens could provide a net benefit to the symbiosis. Polyploidy and heterogeneous haploidy in cultivar fungi is suggested to increase the production and robustness of the symbiosis (Kooij et al., 2015). Diversity in fungal symbionts is also beneficial for other insect-fungal mutualisms in which diverse fungal symbionts are beneficial to the host. Bark beetles, for instance, associate with up to three evolutionarily distinct symbiotic fungi (Bracewell & Six, 2014; Sperling, 2011). This provides a net benefit to the system by allowing the host to adapt over large geographic ranges. Genetic plasticity of arbuscular mycorrhizal fungi (AMF) has enabled them to evade extinction as clonal entities and benefit from associating with many species of trees, sometimes simultaneously (Croll & Sanders, 2009). Long-term clonality of the cultivar fungus of leaf-cutter ants has led an increased risk of infection by a specialized pathogen *Escovopsis* (Ho, 1999; Poulsen & Currie, 2010). This suggests that low amounts of allelic diversity in the fungal cultivar can lead to aggressive infections by pathogens. This also explains why other instances of low fungal diversity – such as colony founding, in which a small inoculum of fungus garden is taken to start a new colony – is associated with extremely high rates of fungus garden failure (which can

approach 84%) (Marti et al., 2015). In all of the above, maintaining high allelic diversity can be associated with improved stability and productivity.

However, our results also suggest that there could be complications associated with the cultivar's genetic state. We observed relatively stable compositions of ASVs in our genetic stability experiment, which suggests that some degree of epistasis occurs between allelic variants resulting in stable allelic combinations. Direct inhibition has occurred between genetically distinct cultivar, making them incompatible for cohabitation (Guo et al., 2017; Poulsen & Boomsma, 2005). This incompatibility negatively influences the productivity of the cultivar, and could feasibly operate on a cellular level. Some cultured isolates from our experiments showed extreme enrichment for comparatively rare cultivar sequence types (WBU in **Figure 4.3.2**). This could be explained as an extreme case of epistatic inhibition of other allelic variants, driving allelic diversity down. If these incompatibilities occurred in already fragile states, such as during colony founding, they could explain the high variability of productivity in newly founded gardens and/or high rates of failure observed (Marti et al., 2015).

Although asexual symbionts are often viewed as beneficial to maintaining stable symbioses (Bronstein et al., 2002), recombination in the asexual symbiont can create variations which can slow rapid adaptations to changing environments in the arbuscular mycorrhizal fungus *Glomus* (Croll & Sanders, 2009). This implies that diverse symbionts are not stable over long periods of time, and certainly not in a rapidly changing environment. However, because all freshly-collected *T. septentrionalis* fungus gardens contained a high level of genetic diversity across the Eastern U.S., the *T. septentrionalis* cultivar may possess a level of allelic diversity that has kept the symbiosis stable since the Pleistocene (Mikheyev et al., 2008). This diversity could be particularly helpful in keeping *T. septentrionalis* colonies adaptable across their wide geographic range; keeping fungus gardens from being exploited by specialist and generalist pathogens; and increasing the odds of successful colony founding by maintaining high levels of

diversity in subcultures. We therefore conclude that this diversity is likely to provide a net benefit or a stable, non-negative influence on the *T. septentrionalis* fungus-growing ant symbiosis.

6 - References

Adams, M. (2017) Characterizing fungal diversity in *Trachymyrmex septentrionalis* fungus gardens (Undergraduate Thesis). University of Connecticut, Storrs, CT.

Albertin, W., & Marullo, P. (2012). Polyploidy in fungi: Evolution after whole-genome duplication. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2497–2509. https://doi.org/10.1098/rspb.2012.0434

Blackwell, M. (2011). The fungi: 1, 2, 3 ... 5.1 million species? *American Journal of Botany*, *98*(3), 426–438. https://doi.org/10.3732/ajb.1000298

Burnett, J. H. (1975) Mycogenetics. London: John Wiley & Sons, Inc.

Burnett, J.H (2003) Fungal populations and species. New York, NY: Oxford University Press Inc.

Bracewell, R. R., & Six, D. L. (2014). Broadscale Specificity in a Bark Beetle – Fungal Symbiosis : a Spatio-temporal Analysis of the Mycangial Fungi of the Western Pine Beetle, 859–870. https://doi.org/10.1007/s00248-014-0449-7

Bronstein, J. L., Rinaldi, S., Law, R., & Gauduchon, M. (2002). Cheating and the evolutionary stability of mutualisms ꜧ ꜧ, (August 2001), 773–780. https://doi.org/10.1098/rspb.2001.1900

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal*, *11*(12), 2639–2643. https://doi.org/10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST + : architecture and applications, BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421

Cooke, R. (1977) The biology of symbiotic fungi. London: John Wiley & Sons Inc.

Cole, G. T. & Kendrick, B. (1981) Biology of conidial fungi. New York, NY: Academic Press Inc.

Croll, D., & Sanders, I. R. (2009). Recombination in Glomus intraradices , a supposed ancient asexual arbuscular mycorrhizal fungus, BMC Evolutionary Biology (9) 13: 1–11. https://doi.org/10.1186/1471-2148-9-13

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data, Microbiome (6)226: 1–14.

de Fine Licht, H. H., Boomsma, J. J., & Tunlid, A. (2014). Symbiotic adaptations in the fungal cultivar of leaf-cutting ants. *Nature Communications*, *5:5675*, 1–10. https://doi.org/10.1038/ncomms6675

Di Franco, A., Poujol, R., Baurain, D., & Philippe, H. (2019). Evaluating the usefulness of alignment filtering methods to reduce the impact of errors on evolutionary inferences. *BMC Evolutionary Biology*, *19*(1), 1–17. https://doi.org/10.1186/s12862-019-1350-2

Edgar, R. C., Drive, R. M., & Valley, M. (2004). MUSCLE : multiple sequence alignment with

high accuracy and high throughput, Nucleic Acids Research *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Guo, C. J., Chang, F. Y., Wyche, T. P., Backus, K. M., Acker, T. M., Funabashi, M., … Fischbach, M. A. (2017). Discovery of Reactive Microbiota-Derived Metabolites that Inhibit Host Proteases. *Cell*, *168*(3), 517–526.e18. https://doi.org/10.1016/j.cell.2016.12.021

Hass, H., & Kerp, H. (1994). Four hundred-million-year-old vesicular arbuscular mycorrhizae, PNAS *91*(December), 11841–11843.

Ho, B. (1999). The agricultural pathology of ant fungus gardens, PNAS Sci *96*(July), 7998–8002.

Howe, J., Schiøtt, M., & Boomsma, J. J. (2018). Horizontal partner exchange does not preclude stable mutualism in fungus-growing ants. *Behavioral Ecology*, In press: 1–11. https://doi.org/10.1093/beheco/ary176

Kerstes, N. A. G., Hoekstra, R. F., Aanen, D. K., Debets, A. J. M., Boomsma, J. J., & de Fine Licht, H. H. (2009). High Symbiont Relatedness Stabilizes Mutualistic Cooperation in Fungus-Growing Termites. *Science*, *326*(5956), 1103–1106. https://doi.org/10.1126/science.1173462

Kooij, P. W., Aanen, D. K., Schiøtt, M., & Boomsma, J. J. (2015). Evolutionarily advanced ant farmers rear polyploid fungal crops. *Journal of Evolutionary Biology*, *28*(11), 1911–1924. https://doi.org/10.1111/jeb.12718

Korb, J., & Aanen, D. K. (2002). The evolution of uniparental transmission of fungal symbionts in fungus-growing termites (Macrotermitinae) in fungus-growing termites (Macrotermitinae), Behav Ecol Sociobiol, 65–71. https://doi.org/10.1007/s00265-002-0559-y

Krishna, K. R. (2005) Mycorrhizas: a molecular analysis. Enfield, NH: Science Publishers Inc.

Lee, K. M., Adams, M., Klassen, J. L. (2018). Evaluation of DESS as a storage medium for microbial community analysis. *PeerJ Preprints*, 27041. https://doi.org/10.7287/peerj.preprints.27041v1

Marti, H. E., Carlson, A. L., Brown, B. V., & Mueller, U. G. (2015). Foundress queen mortality and early colony growth of the leafcutter ant, Atta texana (Formicidae, Hymenoptera). *Insectes Sociaux*, *62*(3), 357–363. https://doi.org/10.1007/s00040-015-0413-7

Mcmurdie, P. J., & Holmes, S. (2013). phyloseq : An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data, PLoS ONE *8*(4): 1-11. https://doi.org/10.1371/journal.pone.0061217

Mikheyev, A. S., Mueller, U. G., & Abbot, P. (2010). Comparative Dating of Attine Ant and Lepiotaceous Cultivar Phylogenies Reveals Coevolutionary Synchrony and Discord, The American Naturalist *175*(6): 126-133. https://doi.org/10.1086/652472

Mikheyev, A. S., Mueller, U. G., & Boomsma, J. J. (2007). Population genetic signatures of diffuse co-evolution between leaf-cutting ants and their cultivar fungi. *Molecular Ecology*, *16*(1), 209–216. https://doi.org/10.1111/j.1365-294X.2006.03134.x

Mikheyev, A. S., Vo, T., & Mueller, U. G. (2008). Phylogeography of post-Pleistocene population expansion in a fungus-gardening ant and its microbial mutualists. *Molecular Ecology*, *17*(20), 4480–4488. https://doi.org/10.1111/j.1365-294X.2008.03940.x

Mueller, U. G., Rehner, S. A., & Schultz, T. R. (1998). The Evolution of Agriculture in Ants Published by : American Association for the Advancement of Science Stable URL : https://www.jstor.org/stable/2895740 REFERENCES Linked references are available on JSTOR for this article : digitize , preserve and extend, *281*(5385), 2034–2038.

Mueller, U. G., Scott, J. J., Ishak, H. D., Cooper, M., & Rodrigues, A. (2010). Monoculture of leafcutter ant gardens. *PLoS ONE*, *5*(9), 1–7. https://doi.org/10.1371/journal.pone.0012668

Muller, H. J. (1964). The relation of recombination to mutual advance, Mutation Research (1), 2–9. https://doi.org/10.1117/12.722789

Nilsson, R. H., Larsson, K., Taylor, A. F. S., Bengtsson-palme, J., Jeppesen, T. S., Schigel, D., Kennedy, P., Picard, K., Glockner, F. O., Tedersoo, L., Saar, I., Koljalg, U., Abarenkov, K. (2019). The UNITE database for molecular identification of fungi : handling dark taxa and parallel taxonomic classifications, Nucleic Acids Research *47*, 259–264. https://doi.org/10.1093/nar/gky1022

Nygaard, S., Hu, H., Li, C., Schiøtt, M., Chen, Z., Yang, Z., … Boomsma, J. J. (2016). Reciprocal genomic evolution in the ant-fungus agricultural symbiosis. *Nature Communications*, *7*, 1–9. https://doi.org/10.1038/ncomms12233

Posada, D. (2008). jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, *25*(7), 1253–1256. https://doi.org/10.1093/molbev/msn083

Poulsen, M., & Boomsma, J. J. (2005). Mutualistic Fungi Control Crop Diversity in Fungus-Growing Ants. *Science*, *307*(5710), 741–744.

Poulsen, M., & Currie, C. R. (2010). Symbiont interactions in a tripartite mutualism: Exploring the presence and impact of antagonism between two fungus-growing ant mutualists. *PLoS ONE*, *5*(1). https://doi.org/10.1371/journal.pone.0008748

Poulsen, M., Currie, C. R., Boomsma, J. J., Poulsen, M., & Al, E. T. (2009). EPHEMERAL WINDOWS OF OPPORTUNITY FOR HORIZONTAL TRANSMISSION OF FUNGAL SYMBIONTS IN LEAF-CUTTING ANTS, Evolution (63) 9: 2235–2247. https://doi.org/10.1111/j.1558-5646.2009.00704.x

Rabeling, C., & Johnson, R. A. (2007). A Review Of The North American Species Of The Fungus-Gardening Ant Genus Trachymyrmex (Hymenoptera: Formicidae), Zootaxa 1664: 1-53. https://doi.org/10.5281/zenodo.180014

R Core Team (2017) R: A Language and Environment for Statistical Computing. https://www.R-project.org/

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewsky, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., Weber, C. F., (2009). Introducing mothur : Open-Source , Platform-Independent , Community-Supported Software for Describing and Comparing Microbial Communities, Applied and Environmental Microbiology *75*(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). HISTORICAL commentary NIH Image to ImageJ : 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., & Levesque, C. A. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode

marker for Fungi, PNAS *109*(16), 6241–6246. https://doi.org/10.1073/pnas.1117018109

Schultz, T. R., & Brady, S. G. (2008). Major evolutionary transitions in ant agriculture. *Proceedings of the National Academy of Sciences*, *105*(14), 5435–5440. https://doi.org/10.1073/pnas.0711024105

Singh, A., & Bhatia, P. (2016). Automated sanger analysis pipeline (ASAP): A tool for rapidly analyzing sanger sequencing data with minimum user interference. *Journal of Biomolecular Techniques*, *27*(4), 129–131. https://doi.org/10.7171/jbt.16-2704-005

Six, D. L. (2007). Dietary Benefits of Fungal Associates to an Eruptive Herbivore : Potential Implications of Multiple Associates on Host, Entomological Society of America 36(6): 1384–1396.

Sperling, F. A. H. (2011). Comparative phylogeography , genetic differentiation and contrasting reproductive modes in three fungal symbionts of a multipartite bark beetle symbiosis, Molecular Ecology (20) 584–600. https://doi.org/10.1111/j.1365-294X.2010.04953.x

Staden R., Beal K.F., Bonfield J.K. (2000) The Staden Package, 1998. In: Misener S., Krawetz S.A. (eds) Bioinformatics Methods and Protocols. Methods in Molecular Biology™, vol 132. Humana Press, Totowa, NJ

Stajich, J. E., Berbee, M. L., Blackwell, M., Hibbett, D. S., James, T. Y., Spatafora, J. W., & Taylor, J. W. (1996). The Fungi. *Current Biology*, *19*(18), 840–845.

Stamatakis, A. (2014). RAxML version 8 : a tool for phylogenetic analysis and post-analysis of large phylogenies. Phylogenetics *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Villa, F., Cappitelli, F., Principi, P., Sorlini, C. (2009) Permeablization method for in-situ investigation of fungal conidia on surfaces. Letters in Applied Microbiology ISSN 0266-8254.

Weber, N. A. (1966). Fungus-growing ants. *Science*, *153*(3736), 587–604.

Weber, N.A. (1972). Gardening ants, the attines. *Science*, 178(4063) 856.

Weiss, S., Van Treuren, W., Lozupone, C., Faust, K., Friedman, J., Deng, Y., Xia, L. C., Xu, Z. Z., Ursell, L., Alm, E. J., Birminham, A., Cram, J., Fuhrman, J. A., Raes, J., Sun, F., Zhou, J., Knight, R. (2016). Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME Journal*, *10*(7), 1669–1681. https://doi.org/10.1038/ismej.2015.235

White, T.J.; Bruns, T.; Lee, S.; Taylor, J., White, T. J., Bruns, T., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols: A Guide to Methods and Applications*, (August 2016), 315–322.

Wilkinson, A., Solan, M., Taylor, A. F. S., Alexander, I. J., & Johnson, D. (2010). Intraspecific Diversity Regulates Fungal Productivity and Respiration. PLoS ONE *5*(9), 1–9. https://doi.org/10.1371/journal.pone.0012604

Zeigler, R. S., Scott, R. P., Leung, H., Bordeos, A. A., Kumar, J., & Nelson, R. J. (1996). Evidence of Parasexual Exchange of DNA in the Rice Blast Fungus Challenges Its Exclusive Clonality. Phytopathology 87:284-294.

7 – Appendix

7.1 – Figures

Table 4.2.1 Sample collection metadata of freshly-collected fungus gardens

Sample_ID	State	Colony_or_Host_ID	Collection Month	Collection Year	GPS_Latitude	GPS_Longitude	Geographic_Location	Site
JKA000222	NJ	JKH000038	June	2014	39.55076N	74.31335W	Brendon T. Byrne State Forest	WSF-G
JKA000263	NJ	JKH000043	June	2014	39.42319N	74.39340W	Brendon T. Byrne State Forest	WSF-H
JKA000344	NJ	JKH000055	June	2014	39.466871N	74.37842W	Wharton State Forest	WSF-H
JKA000348	NJ	JKH000056	June	2014	39.46684N	74.37833W	Wharton State Forest	WSF-H
JKA000355	NJ	JKH000057	June	2014	39.4646693 N	74.37835W	Wharton State Forest	WSF-H
JKA000507	NJ	JKH000063	July	2014	40.53558	72.49347	Robert Murphy County Park	WSF-H
JKA000520	NJ	JKH000065	July	2014	39.52365N	74.31294W	Brendon T. Byrne State Forest	WSF-H
JKA000547	NJ	JKH000067	July	2014	39.52381N	74.31310W	Brendon T. Byrne State Forest	WSF-M
JKA000558	NJ	JKH000068	July	2014	39.52320N	74.31244W	Brendon T. Byrne State Forest	WSF-Q
JKA000594	NJ	JKH000072	July	2014	39.52353N	74.31199W	Wharton State Forest	WSF-Q
JKA000597	NJ	JKH000072	July	2014	39.52353N	74.31199W	Wharton State Forest	WSF-Q
JKA000602	NJ	JKH000073	July	2014	39.42730N	74.33810W	Wharton State Forest	WSF-Q
JKA000610	NJ	JKH000073	July	2014	39.42730N	74.33810W	Wharton State Forest	WSF-Q
JKA000614	NJ	JKH000074	July	2014	39.42733N	74.33869W	Wharton State Forest	WSF-B
JKA000625	NJ	JKH000074	July	2014	39.42733N	74.33869W	Wharton State Forest	WSF-B
JKA000628	NJ	JKH000075	July	2014	39.42752N	74.33875W	Wharton State Forest	WSF-B
JKA000631	NJ	JKH000076	July	2014	39.42748N	74.33861W	Wharton State Forest	WSF-B
JKA000641	NJ	JKH000077	July	2014	39.42747N	74.33861W	Wharton State Forest	WSF
JKA000642	NJ	JKH000077	July	2014	39.42747N	74.33861W	Wharton State Forest	WSF
JKA000650	NJ	JKH000078	July	2014	39.42749N	73.33864W	Wharton State Forest	WSF

JKA000660	NJ	JKH000079	July	2014	39.42765	74.33865W	Wharton State Forest	WSF
JKA000827	FL	JKH000099	November	2014	28.71057	81.48445	Wekiwa Springs	WSF
JKA000906	FL	JKH000109	November	2014	28.57105	82.25468	Withalacochee	WSF
JKA001085	GA	JKH000128 Bottom	May	2015	NA	NA	George L. Smith State Park	WSF
JKA001100	GA	JKH000128 Top	May	2015	NA	NA	George L. Smith State Park	WSF
JKA001123	GA	JKH000129 Top	May	2015	NA	NA	George L. Smith State Park	WSF
JKA001126	GA	JKH000130	May	2015	NA	NA	George L. Smith State Park	WSF
JKA001150	GA	JKH000131 Bottom	May	2015	NA	NA	George L. Smith State Park	WSF
JKA001165	GA	JKH000133	May	2015	NA	NA	George L. Smith State Park	WSF
JKA001174	GA	JKH000134 Bottom	May	2015	NA	NA	Yuchi Wildlife Management Area	WSF
JKA001204	GA	JKH000134 Top	May	2015	NA	NA	Yuchi Wildlife Management Area	WSF
JKA001184	GA	JKH000135 Top	May	2015	NA	NA	Yuchi Wildlife Management Area	WSF
JKA001227	GA	JKH000136 Bottom	May	2015	NA	NA	Yuchi Wildlife Management Area	BTB-B
JKA001243	GA	JKH000137 Top	May	2015	NA	NA	Yuchi Wildlife Management Area	BTB-B
JKA001294	GA	JKH000141	May	2015	32.88268	-81.95129	Magnolia Springs State Park	BTB-B
JKA001308	GA	JKH000142 Bottom	May	2015	32.88267	-81.95127	Magnolia Springs State Park	BTB-B
JKA001409	LA	JKH000149 Bottom	May	2015	33.01135	-81.89949	Alexander Wildlife Management Area	BTB-B
JKA001822	NC	JKH000155	June	2015	35.86074	-78.76111	William B Umpstead State Park	BTB-B
JKA001486	NC	JKH000156	June	2015	34.68159	-78.59554	William B Umpstead State Park	BTB
JKA001544	NC	JKH000163	June	2015	34.58159	-78.44874	William B Umpstead State Park	BTB
JKA001565	NC	JKH000164	June	2015	34.5814	-78.44859	William B Umpstead State Park	BTB
JKA001596	NC	JKH000165	June	2015	34.40608	-78.98442	William B Umpstead State Park	BTB
JKA001620	NC	JKH000168	June	2015	34.38903	-79.00242	William B Umpstead State Park	BTB
JKA001704	NC	JKH000172	June	2015	34.91705	-79.35333	William B Umpstead State Park	BTB-M
JKA001725	NC	JKH000174	June	2015	34.91711	-79.35344	William B Umpstead State Park	BTB-M
JKA002341	NY	JKH000181	July	2015	40.8926	-72.822483	Robert Cushman Murphy County Park	BTB-M
JKA002735	LA	JKH000185	May	2016	31.1375	-92.484961	Alexander Forest Wildlife Mangement Area	GLS

JKA002747	LA	JKH000186	May	2016	31.12922	-92.49341	Alexander Forest Wildlife Mangement Area	GLS
JKA002762	LA	JKH000187	May	2016	31.0513	-93.4027	Clear Creek Wildlife Management Area	GLS
JKA002778	LA	JKH000188	May	2016	31.05115	-93.40287	Clear Creek Wildlife Management Area	GLS
JKA002832	LA	JKH000192	May	2016	31.01988	-93.06712	Fort Polk	GLS
JKA002875	LA	JKH000195	May	2016	31.01987	-93.06717	Fort Polk	GLS
JKA002896	LA	JKH000197	May	2016	31.11343	-92.4688	Alexander Forest Wildlife Mangement Area	YWM
JKA002913	LA	JKH000198	May	2016	31.1134	-92.46855	Alexander Forest Wildlife Mangement Area	YWM
JKA002925	LA	JKH000199	May	2016	31.11346	-92.46857	Alexander Forest Wildlife Mangement Area	YWM
JKA003000	FL	JKH000204	June	2016	30.00193	-84.53964	Tate's Hell State Forest	YWM
JKA003013	FL	JKH000206	June	2016	NA	NA	Lake Talquin State Forest:Fort Braden tract	YWM
JKA003025	FL	JKH000207	June	2016	30.4401	-84.49546	Lake Talquin State Forest:Fort Braden tract	MSS
JKA003060	FL	JKH000210	June	2016	30.43928	-84.49528	Lake Talquin State Forest:Fort Braden tract	MSS
JKA003083	FL	JKH000212	June	2016	30.43821	-84.49589	Lake Talquin State Forest:Fort Braden tract	WBU
JKA003095	FL	JKH000213	June	2016	30.46497	-84.36418	Lake Talquin State Forest: South Ochlockonee WMA	WBU
JKA003113	FL	JKH000215	June	2016	30.47538	-84.38263	Lake Talquin State Forest: South Ochlockonee WMA	WBU
JKA003144	FL	JKH000219	June	2016	30.34233	-84.25484	Wakulla State Forest:Woodville tract	WBU
JKA003156	FL	JKH000221	June	2016	30.3425	-84.2549	Wakulla State Forest:Woodville tract	WBU
JKA003196	FL	JKH000226	June	2016	30.34236	-84.25365	Wakulla State Forest:Woodville tract	WBU
JKA003440	NY	JKH000230	July	2017	40.89286	-72.82242	Robert Cushman Murphy Park/Otis Pike Preserve	WBU
JKA003463	NY	JKH000233	July	2017	39.87339	-74.52663	Brendan T Byrne State Forest:Brendan T Byrne Campground	WBU
JKA003473	NY	JKH000232	July	2017	39.87326	-74.52628	Brendan T Byrne State Forest:Brendan T Byrne Campground	AFM
JKA003485	NY	JKH000234	July	2017	39.87341	-74.52663	Brendan T Byrne State Forest:Brendan T Byrne Campground	AFM
JKA003497	NY	JKH000235	July	2017	39.87261	-74.52644	Brendan T Byrne State Forest:Brendan T Byrne Campground	AFM
JKA003501	NY	JKH000235	July	2017	39.87261	-74.52644	Brendan T Byrne State Forest:Brendan T Byrne Campground	AFM
JKA003520	NY	JKH000237	July	2017	39.87234	-74.52605	Brendan T Byrne State Forest:Brendan T Byrne Campground	AFM

JKA003559	NY	JKH000240	July	2017	39.91817	-74.52167	USA;New Jersey;Brendan T Byrne State Forest;Mount Misery	AFM
JKA003571	NY	JKH000241	July	2017	39.9179	-74.52129	Brendan T Byrne State Forest :Mount Misery	CWC
JKA003581	NY	JKH000242	July	2017	39.91759	-74.52072	Brendan T Byrne State Forest :Mount Misery	CWC
JKA003591	NY	JKH000243	July	2017	39.688	-74.54983	Wharton State Forest :Godfrey Bridge Family Camp	FP
JKA003607	NY	JKH000244	July	2017	39.711	-74.56396	Wharton State Forest:Hawkin Bridge Camp	FP
JKA003619	NY	JKH000245	July	2017	39.71257	-74.56434	Wharton State Forest:Hawkin Bridge Camp	WS
JKA003634	NY	JKH000246	July	2017	39.71121	-74.56368	Wharton State Forest:Hawkin Bridge Camp	Wit
JKA003645	NY	JKH000248	July	2017	39.711	-74.564	Wharton State Forest:Hawkin Bridge Camp	THS
JKA003656	NY	JKH000249	July	2017	39.71074	-74.56433	Wharton State Forest:Hawkin Bridge Camp	LTF-B
JKA003662	NY	JKH000247	July	2017	39.71252	-74.56438	Wharton State Forest:Hawkin Bridge Camp	LTF-B
JKA003712	NY	JKH000253	July	2017	39.722663	-74.69146	Wharton State Forest:Road to Mullica Camp	LTF-B
JKA003741	NY	JKH000256	July	2017	39.70969	-74.66375	Wharton State Forest:Quaker Bridge	LTF-B
JKA003766	NY	JKH000258	July	2017	39.70916	-74.66415	Wharton State Forest:Quaker Bridge	LTF-S
JKA003790	NY	JKH000260	July	2017	39.70961	-74.66337	Wharton State Forest:Quaker Bridge	LTF-S
JKA003815	NY	JKH000262	July	2017	39.70931	-74.66341	Wharton State Forest:Quaker Bridge	WSF-W
JKA003827	NY	JKH000263	July	2017	39.7096	-74.6639	Wharton State Forest:Quaker Bridge	WSF-W
JKA003836	NY	JKH000264	July	2017	39.77803	-74.63095	Wharton State Forest:Batona Camp	WSF-W
JKA003852	NY	JKH000265	July	2017	39.77801	-74.63147	Wharton State Forest:Batona Camp	RMC
JKA003864	NY	JKH000266	July	2017	39.77818	-74.6315	Wharton State Forest:Batona Camp	RCM-O
JKA003878	NY	JKH000267	July	2017	39.77798	-74.63104	Wharton State Forest:Batona Camp	RCM-O

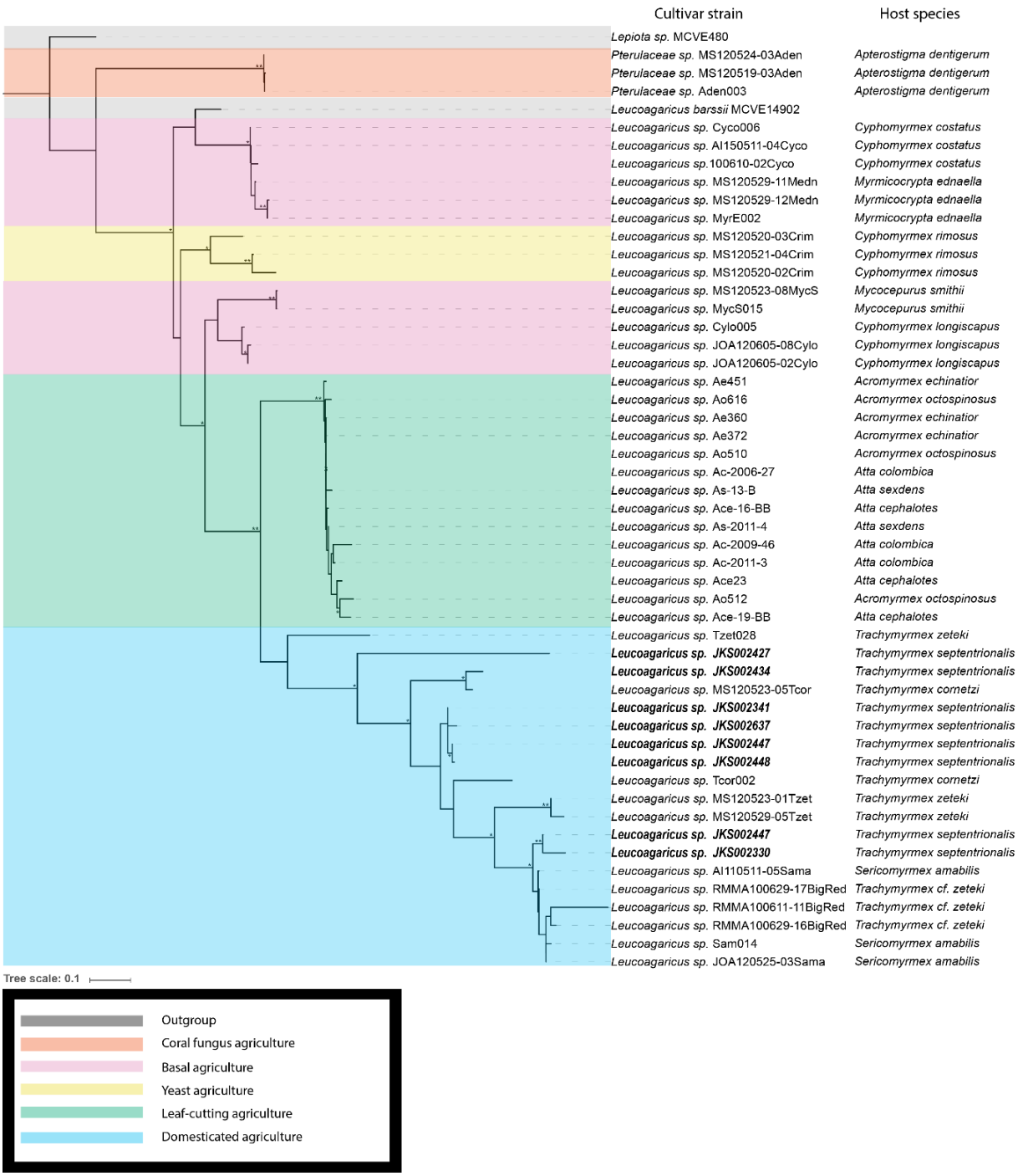

Figure 4.1.1 – Phylogeny of attine agriculture. Maximum likelihood phylogeny of cultivar fungus ITS sequences in different attine agriculture types. Colors represent types of agriculture; outgroups are labeled in gray. Cultivar strain names (left column) are paired with the attine host species (right column) they were isolated from. Sequences from this study are in bold. Nodes labeled with * and ** have posterior probabilities of ≥65 < 90 and ≥ 90 ≤ 100, respectively. The scale bar corresponds to 0.1 substitutions per site.

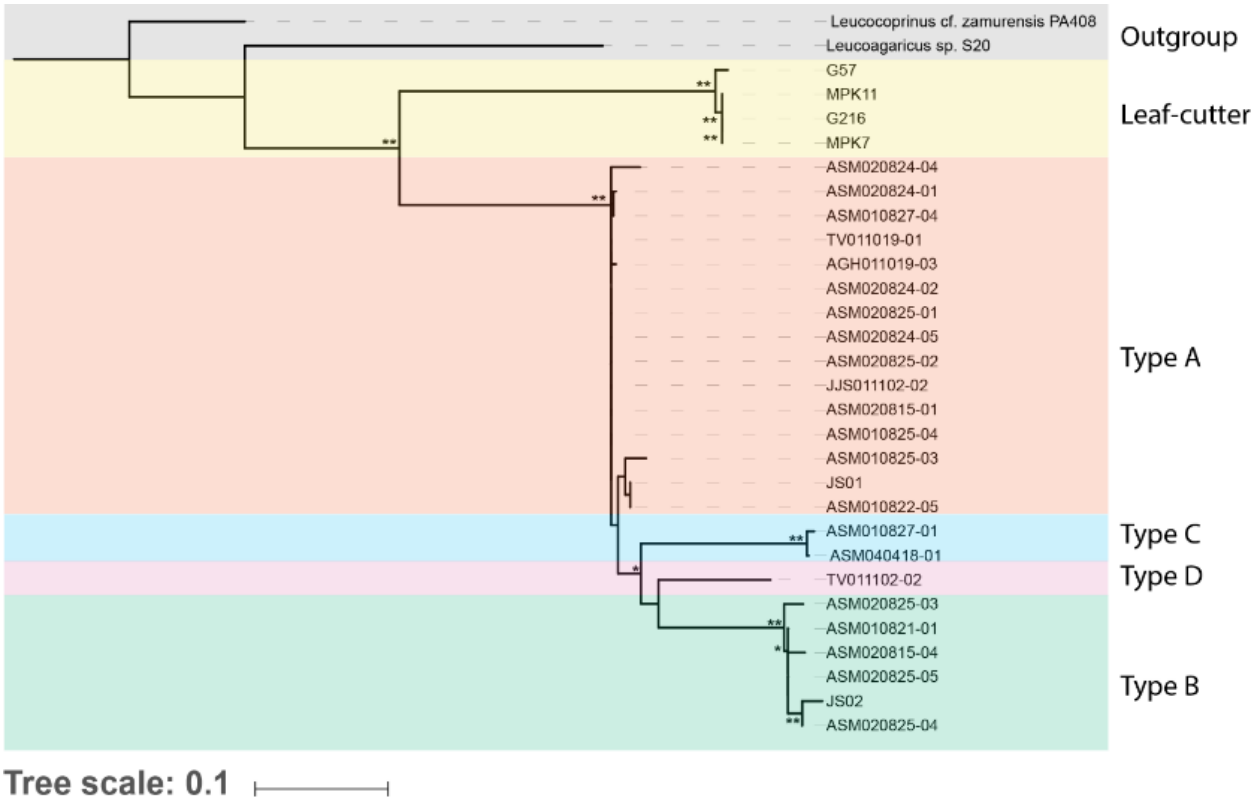

Figure 4.2.1 – ITS phylogeny of cultivar isolates from Mikheyev et al. 2008. Maximum likelihood phylogeny of *Leucoagaricus* ITS sequences. Colors over tree distinguish cultivar types as published in Mikheyev et al 2008. Nodes labeled with * and ** have posterior probabilities of ≥65 < 90 and ≥ 90 ≤ 100, respectively. The scale bar corresponds to 0.1 substitutions per site.

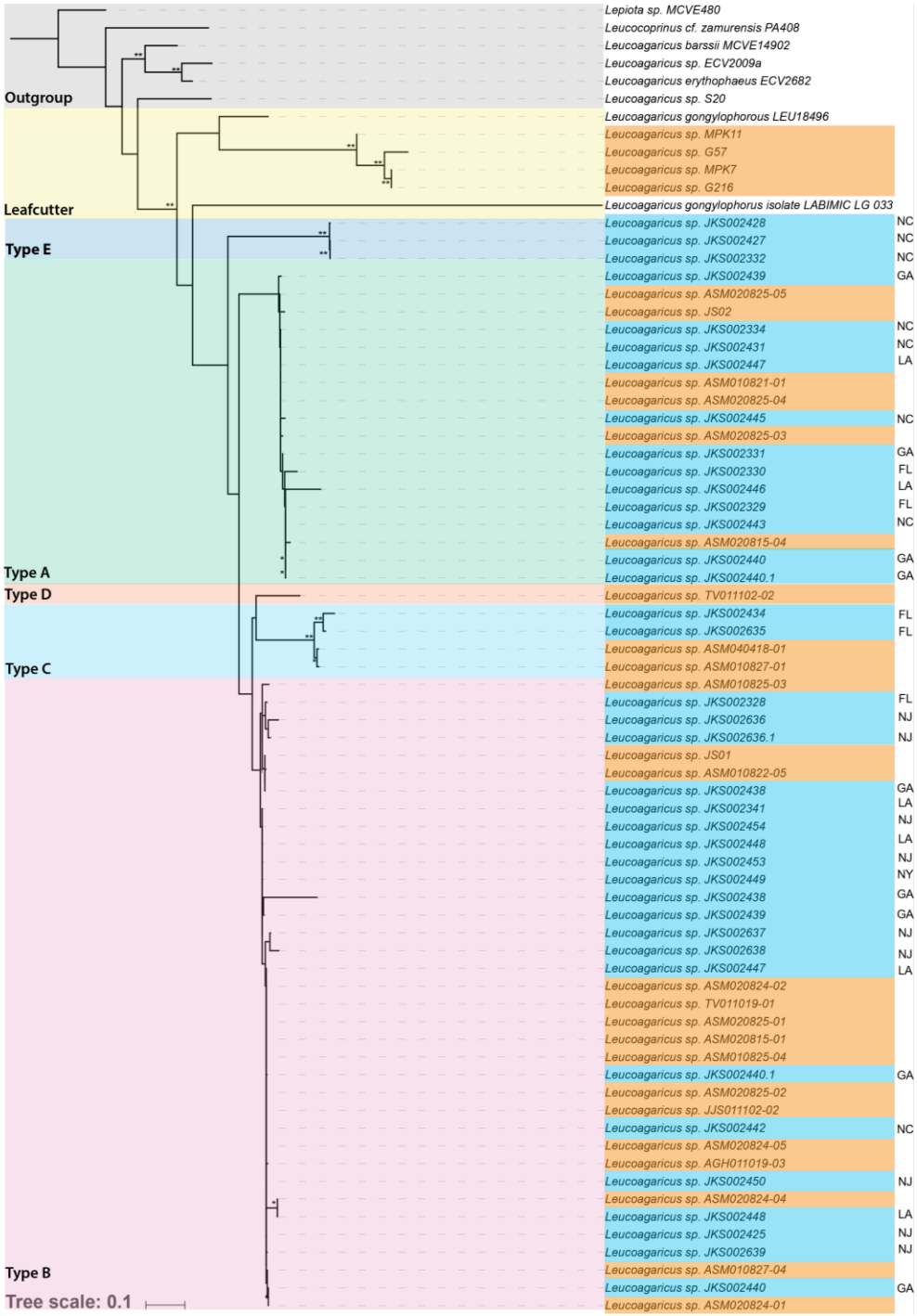

Figure 4.1.3 – Phylogeny of cultured cultivar isolates reared by *T. septentrionalis.* Maximum likelihood phylogeny of *Leucoagaricus* cultivar ITS sequences reared by *T. septentrionalis* fungus-growing ants. Colors over tree distinguish cultivar types. Strain names are color-coded by study of origin: Orange – Mikheyev et al 2008; Blue – this work; No color – other. State of origin appears in the far right column. Nodes labeled with * and ** have posterior probabilities of ≥65 < 90 and ≥ 90 ≤ 100, respectively. The scale bar corresponds to 0.1 substitution per site.

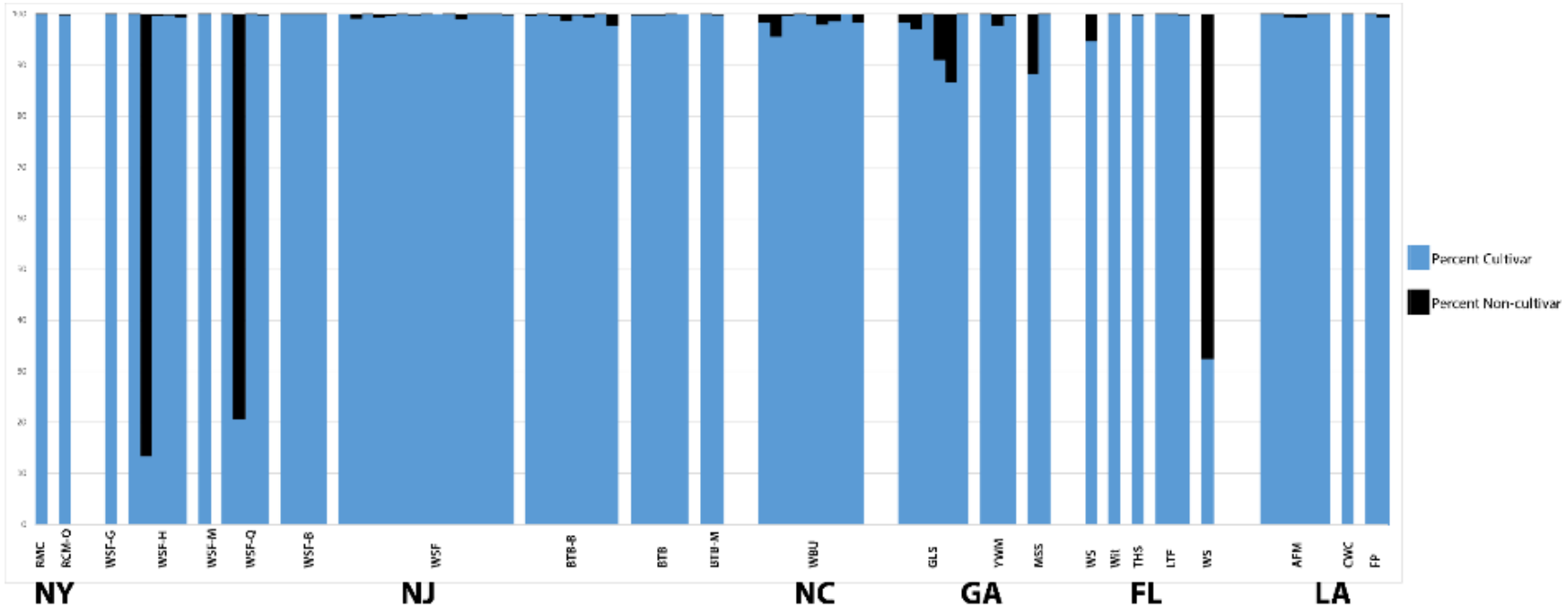

Figure 4.2.1 – Percent abundance of cultivar and non-cultivar ASVs in freshly-collected *T. septentrionalis* fungus gardens. Labels beneath indicate state and collection site of origin, following Table 2.1.

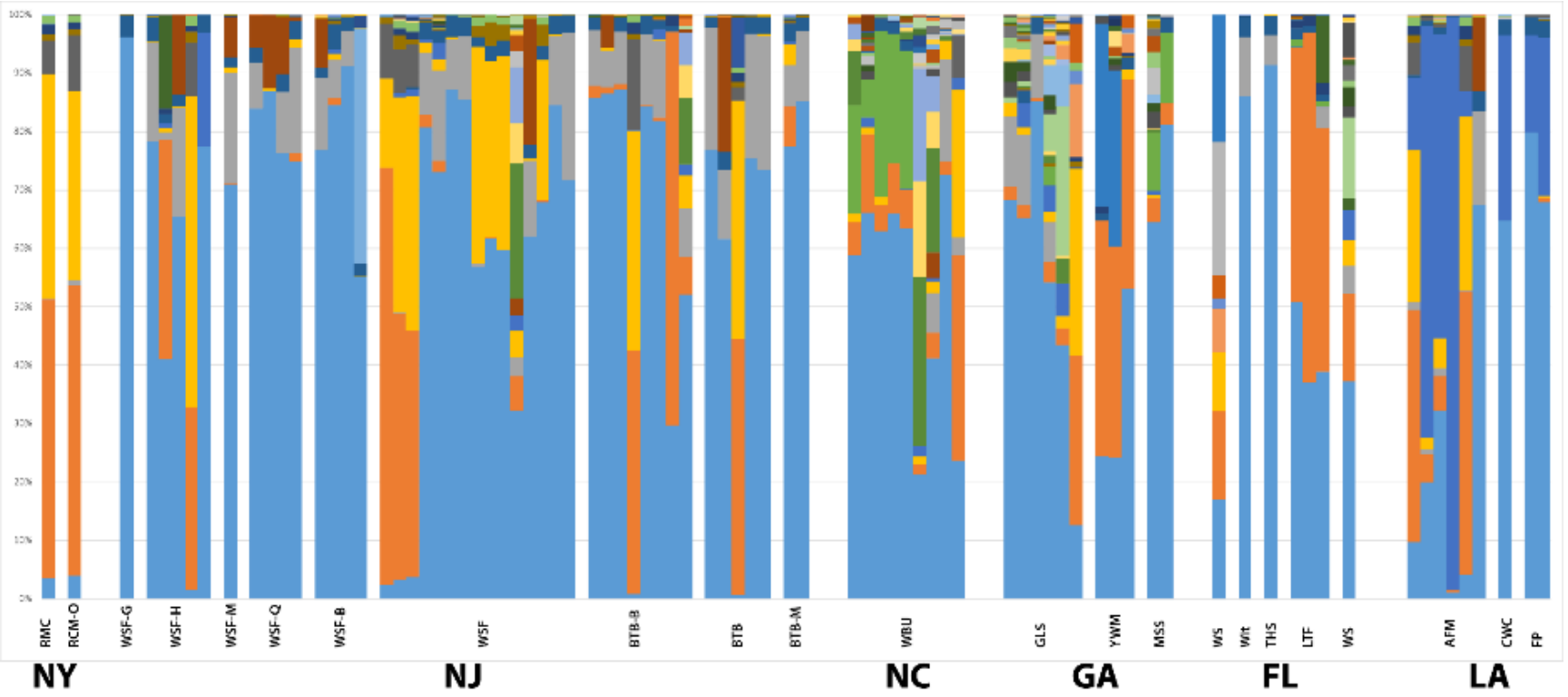

Figure 4.2.2 – Percent abundance of cultivar ASVs in freshly-collected *T. septentrionalis* fungus gardens. Labels beneath indicate state and collection site of origin following Table 2.1. Each unique color represents a different cultivar ASV. Non-cultivar fungi were removed.

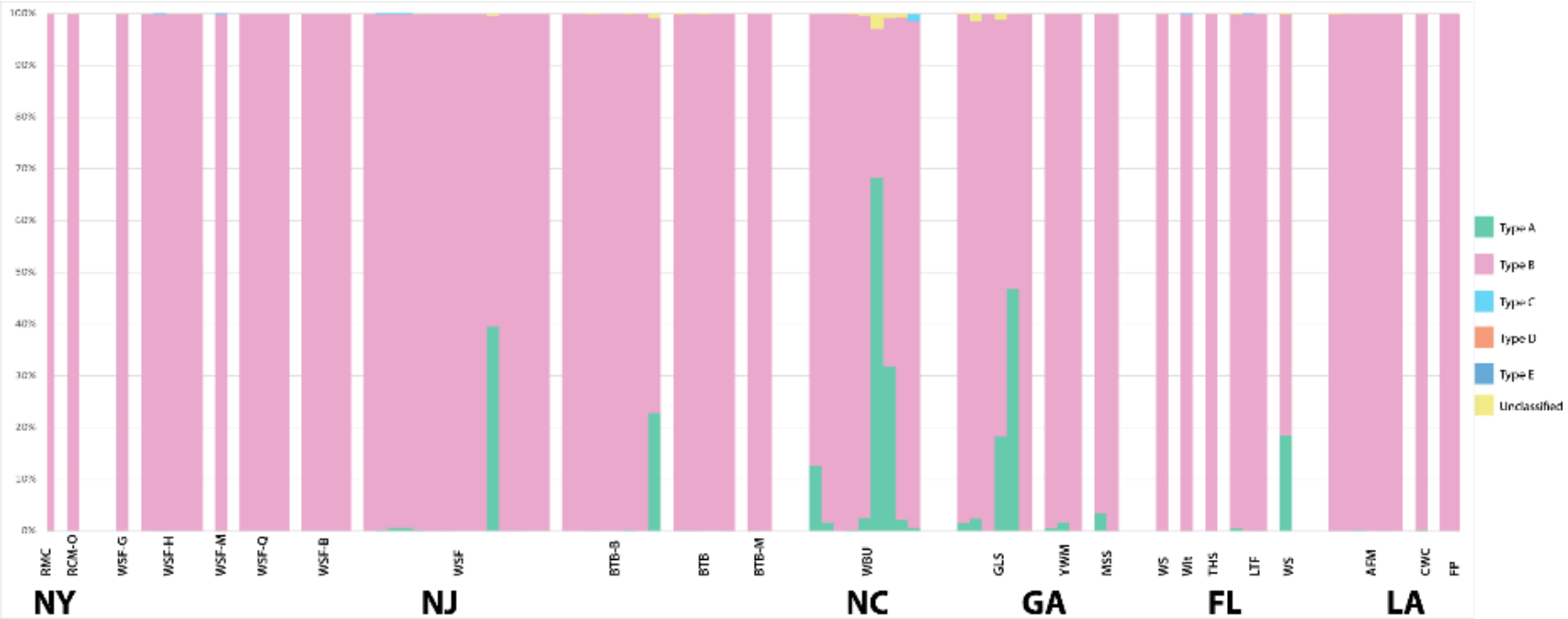

Figure 4.2.4 – Percent abundance of cultivar sequence types in freshly-collected cultivar ASVs. Labels beneath indicate state and collection site of origin following Table 2.1. Colors correspond to cultivar types established in Figure 4.1.3.

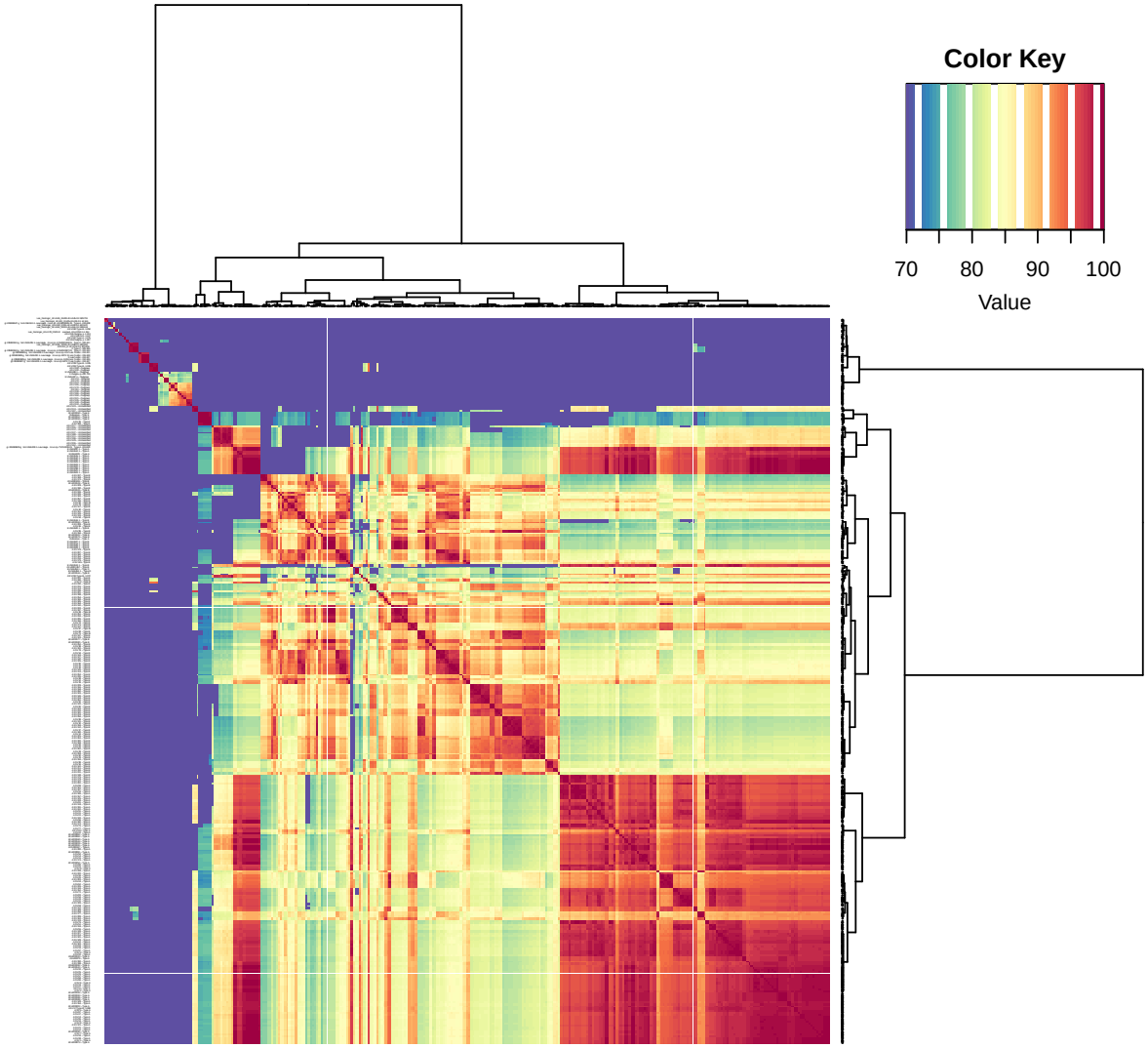

Figure 4.2.3 – Heatmap of BLAST comparison between cultivar ASVs and reference sequences from Figure 4.1.3. ITS ASV and reference sequences are labelled on the left side. The distance between alignment values as calculated using the Euclidean method and clustered using Ward's method.

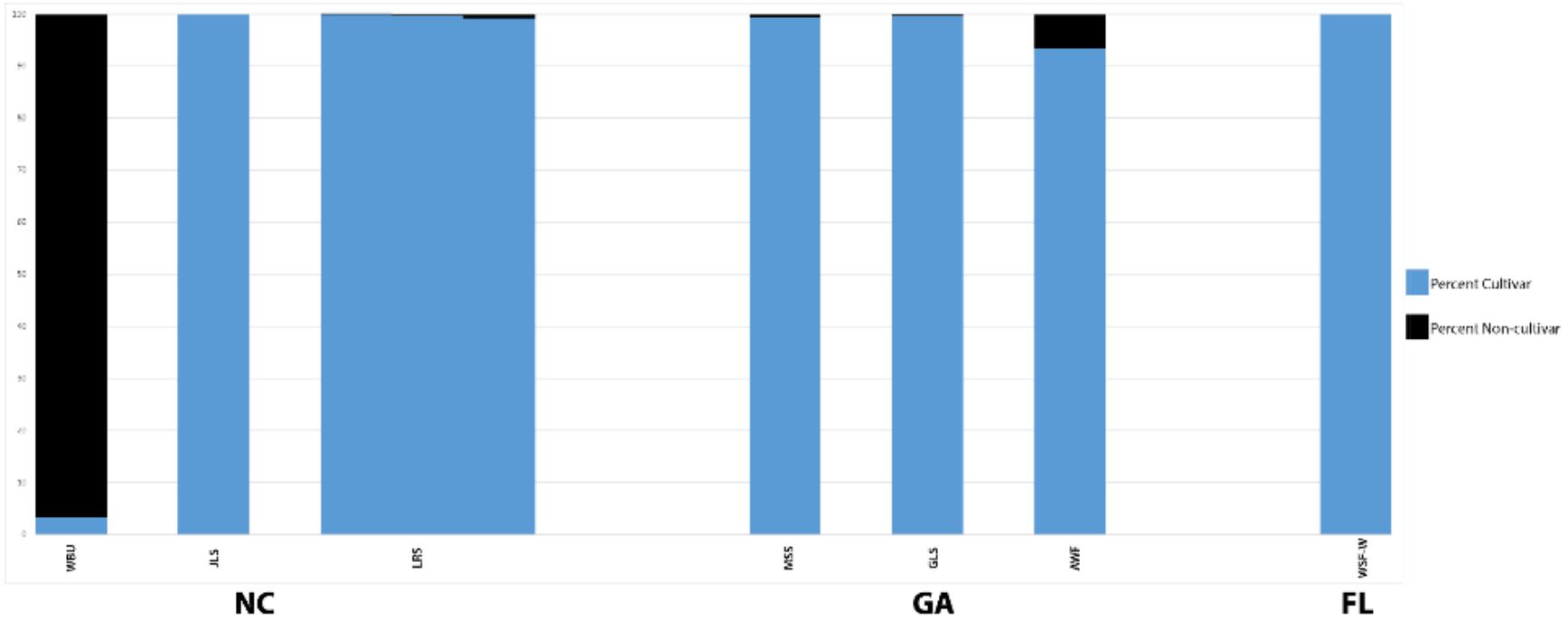

Figure 4.3.1 – Percent abundance of cultivar and non-cultivar ASVs in *T. septentrionalis* cultured isolates. Labels beneath indicate state and collection site of origin following Table 2.1

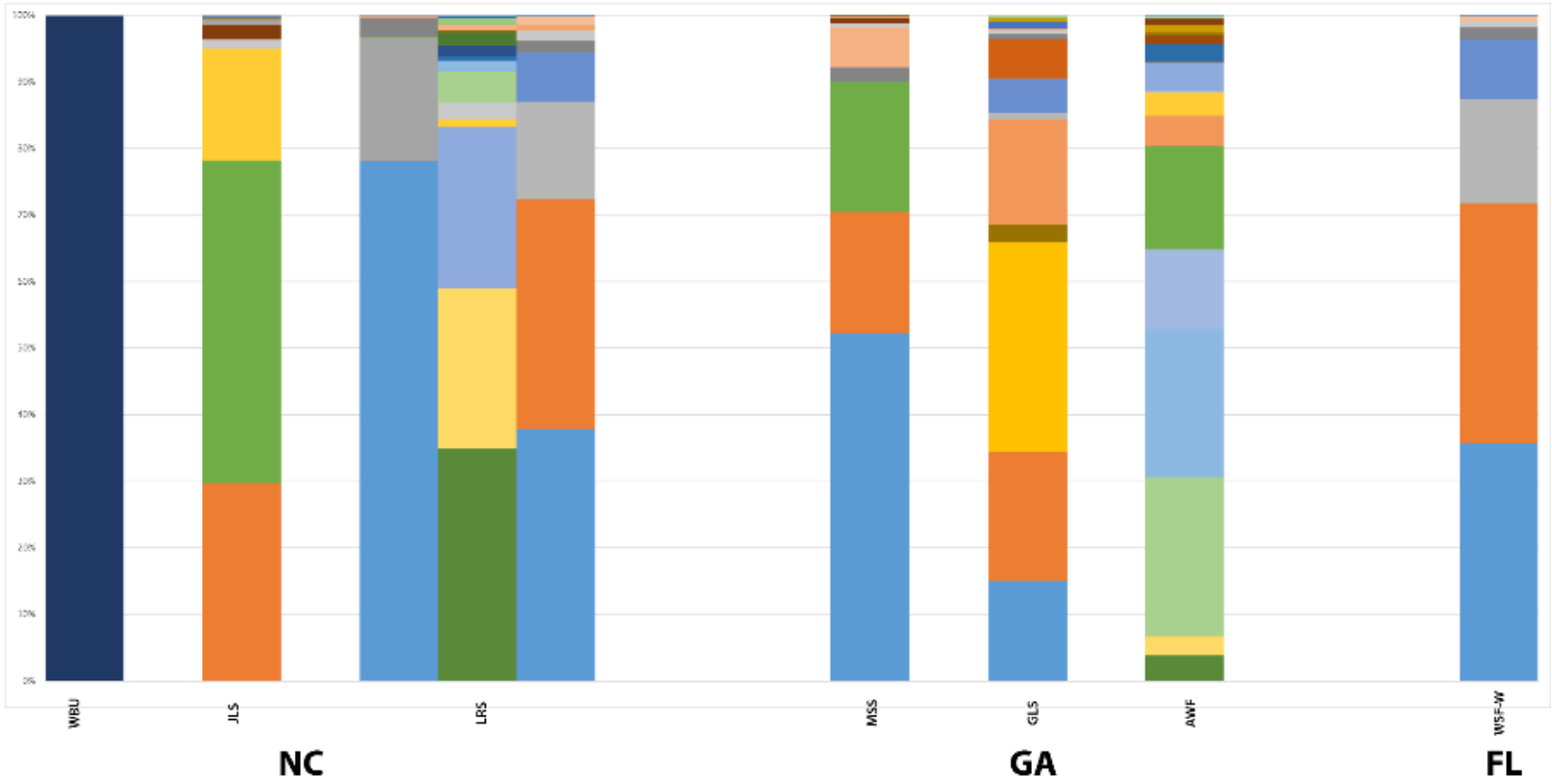

Figure 4.3.2 – Percent abundance of cultivar ASVs in *T. septentrionalis* cultured isolates. Labels beneath indicate state and collection site of origin, following Table 2.1. Each color represents a unique cultivar ASV.

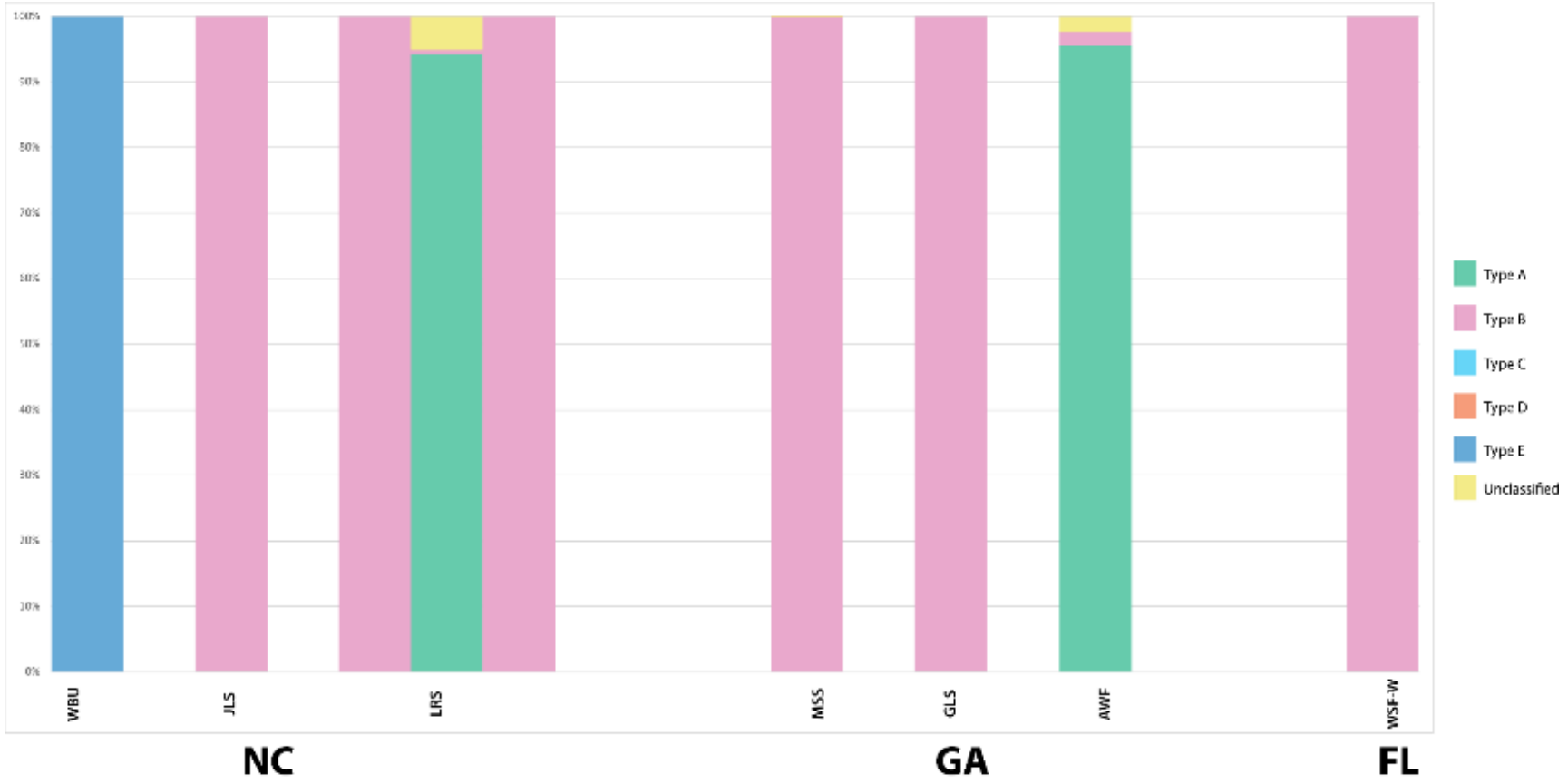

Figure 4.3.3 – Percent abundance of cultivar sequence types in cultured isolate cultivar ASVs. Labels beneath indicate state and collection site of origin following Table 2.1. Colors correspond to cultivar types established in Figure 4.1.3.

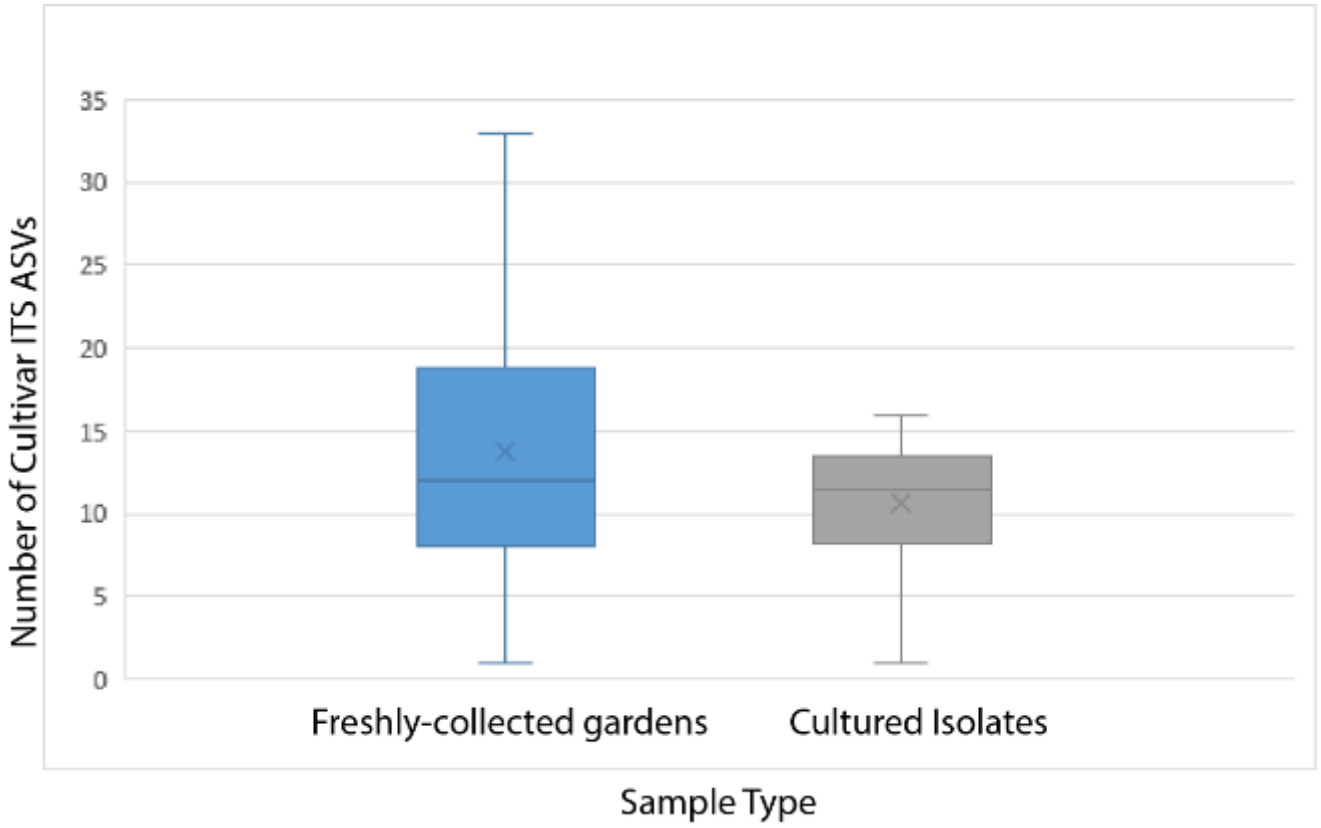

Figure 4.3.4 – Cultivar ASV alpha diversity comparison between freshly-collected gardens and cultured isolates. Total number of cultivar ASVs in freshly-collected gardens and cultured isolates. Plots show mean, first and third quartiles, and maximum/ minimum values.

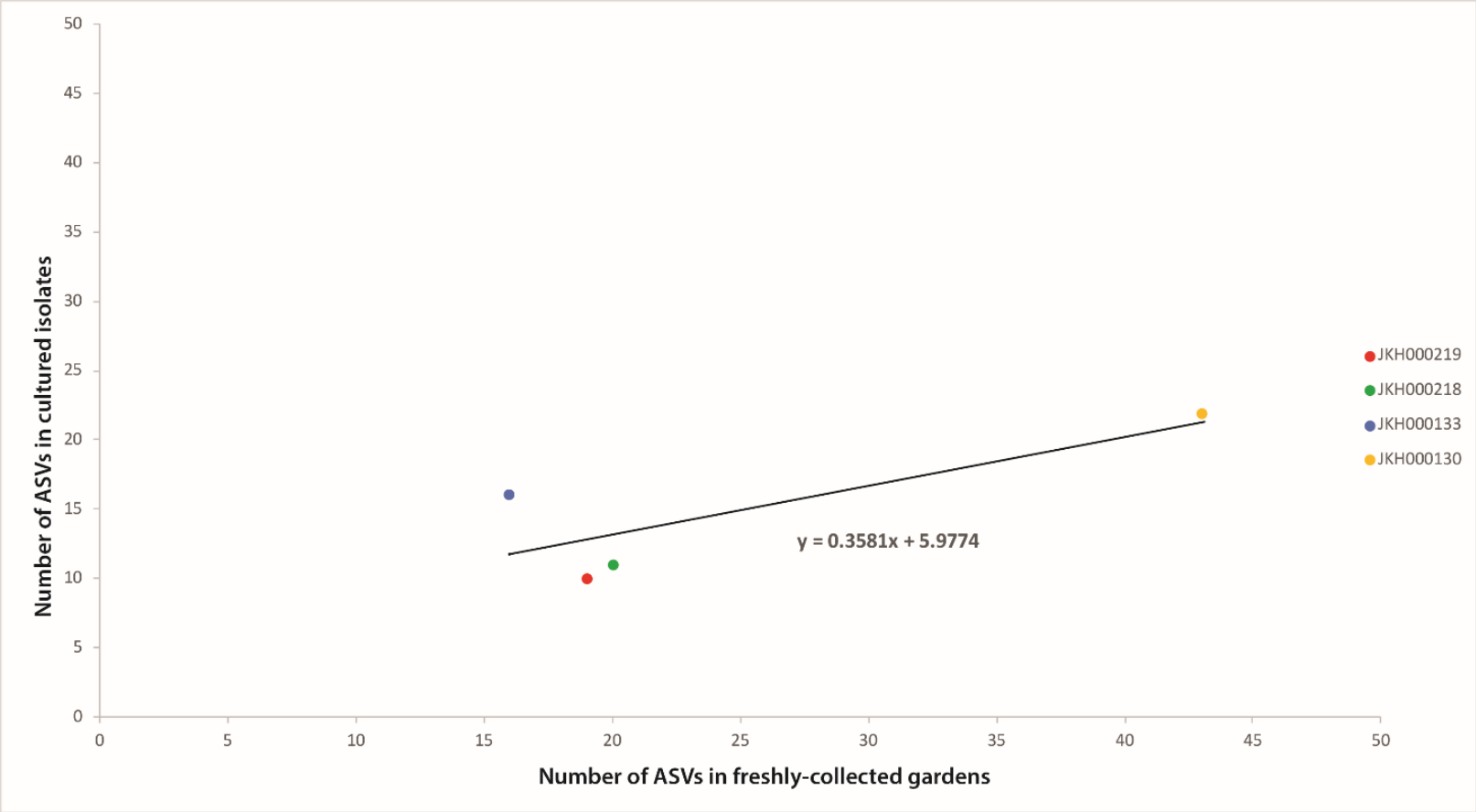

Figure 4.3.5 – Correlation between the alpha diversity of freshly-collected gardens and their corresponding cultured isolates. Point color corresponds to the wild garden colony ID. The plotted linear relationship of points has an equation of y = 0.3581x + 5.9774.

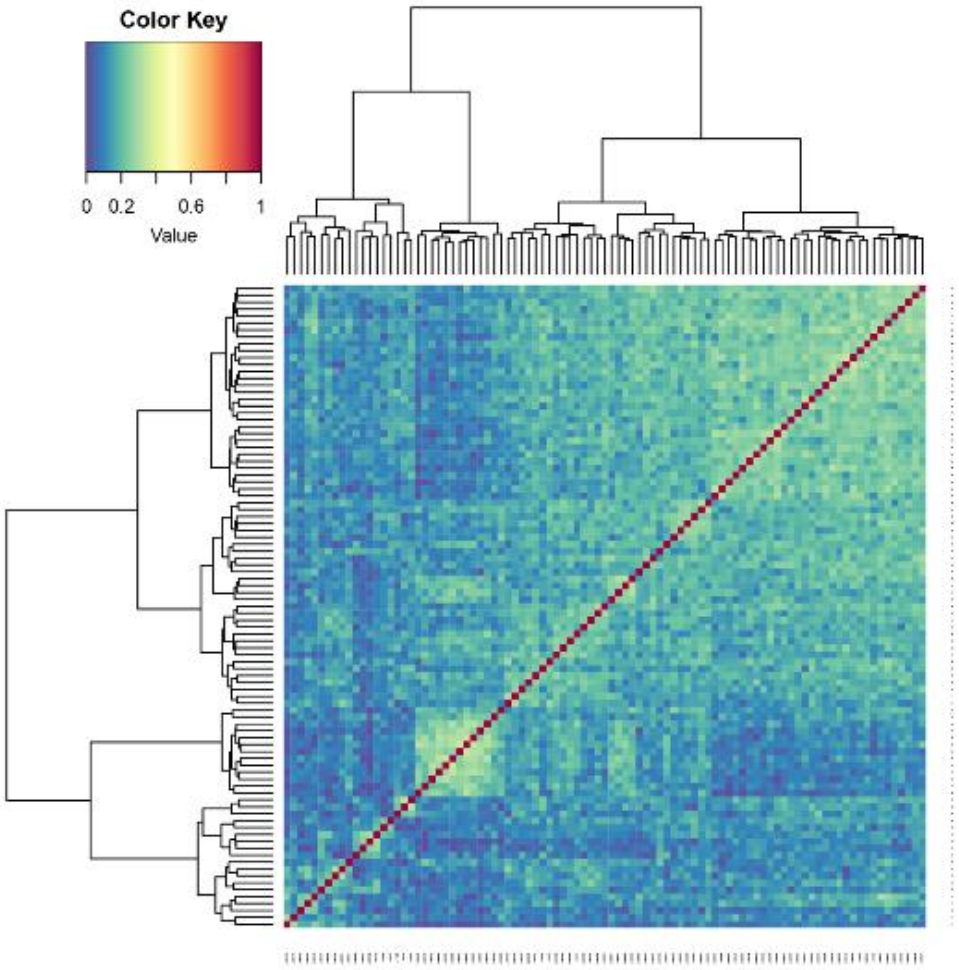

Figure 4.4.1 – Heatmap of co-occurrence between cultivar ASVs in freshly-collected *T. septentrionalis* fungus gardens. Cultivar ASVs are labelled on the left side. Correlation coefficients were calculated based on SparCC's correlation method and clustering was calculated using Ward's method.

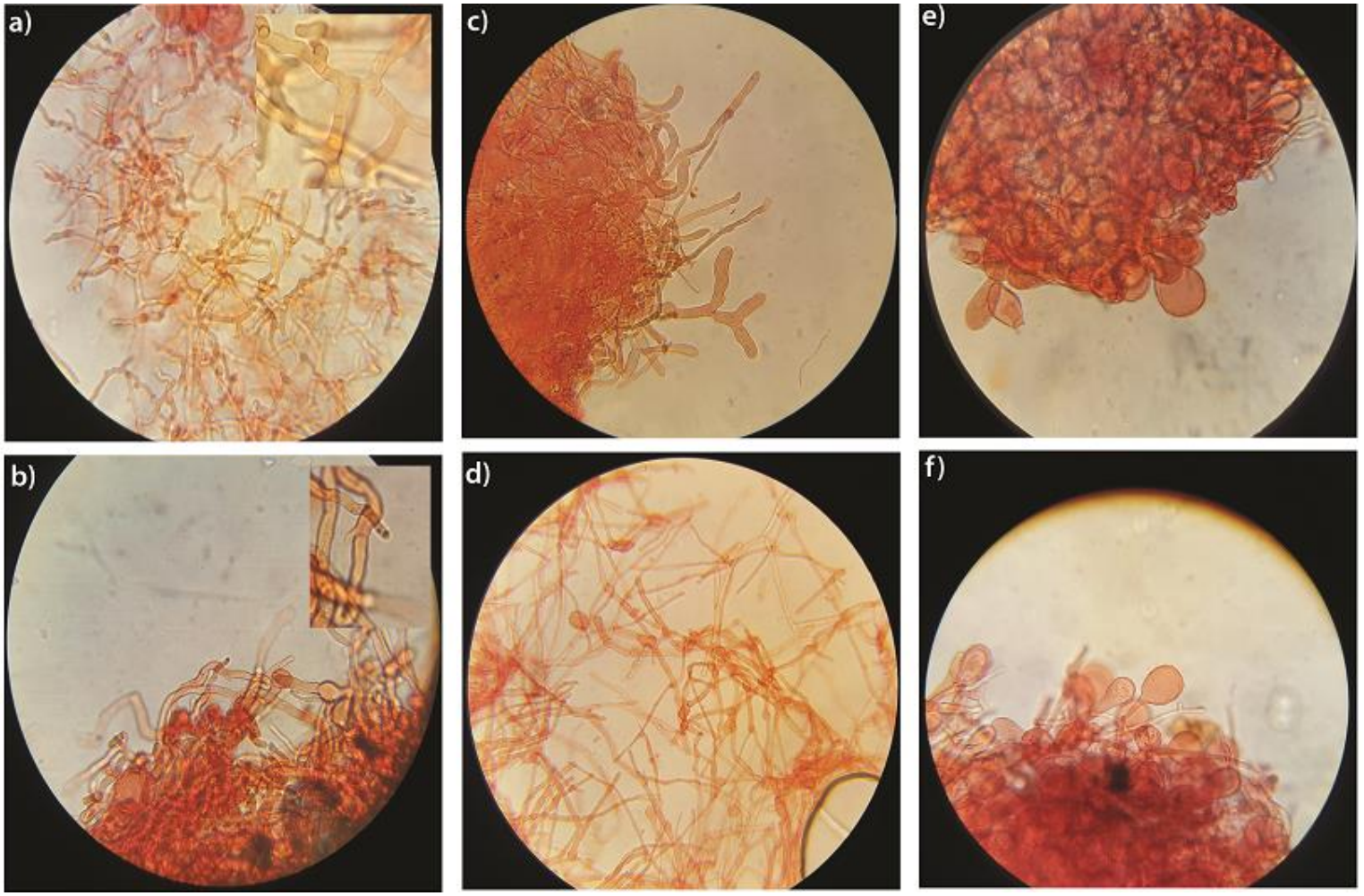

Figure 4.5.1 – Histology of cultured *T. septentrionalis* isolates. Comparisons of cell types and morphological differences between cultured isolates stained with 0.1% congo red. Vegetative hyphae (a,b) with vignettes showing anasmotoses and clamp connections, respectively. Clavulate gongylidia (c) and non-terminal gongylidia (d) compared to typical tightly-packed and globose gongylidia (e,f).

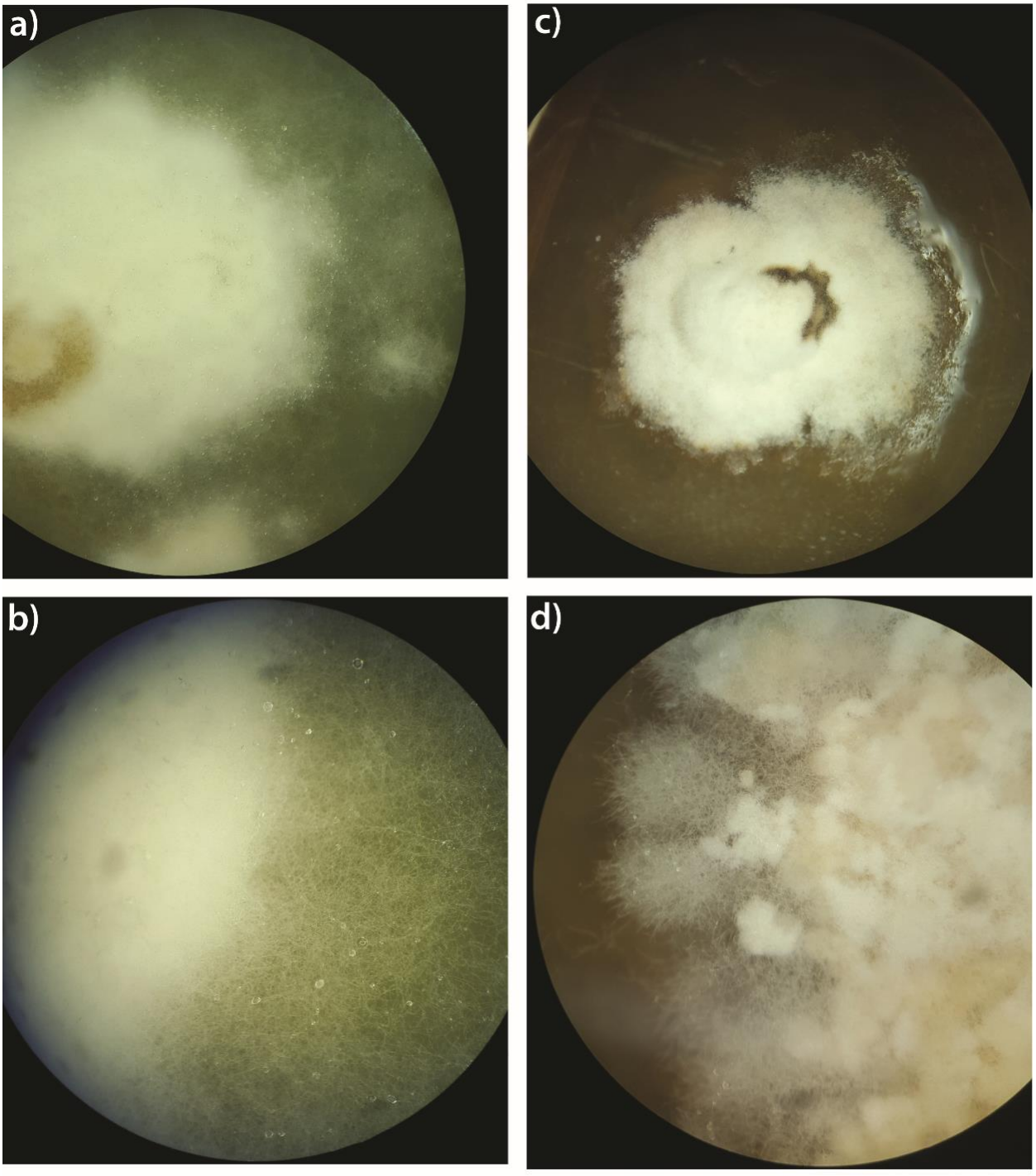

Figure 4.5.2 – Macroscopic morphology of cultured *T. septentrionalis* cultivar isolates. Cultured isolate morphology grown on PDA. Indistinct staphylae and irregular hyphae (a,b) compared to normal colony morphology (c) and dense staphylae production (d).

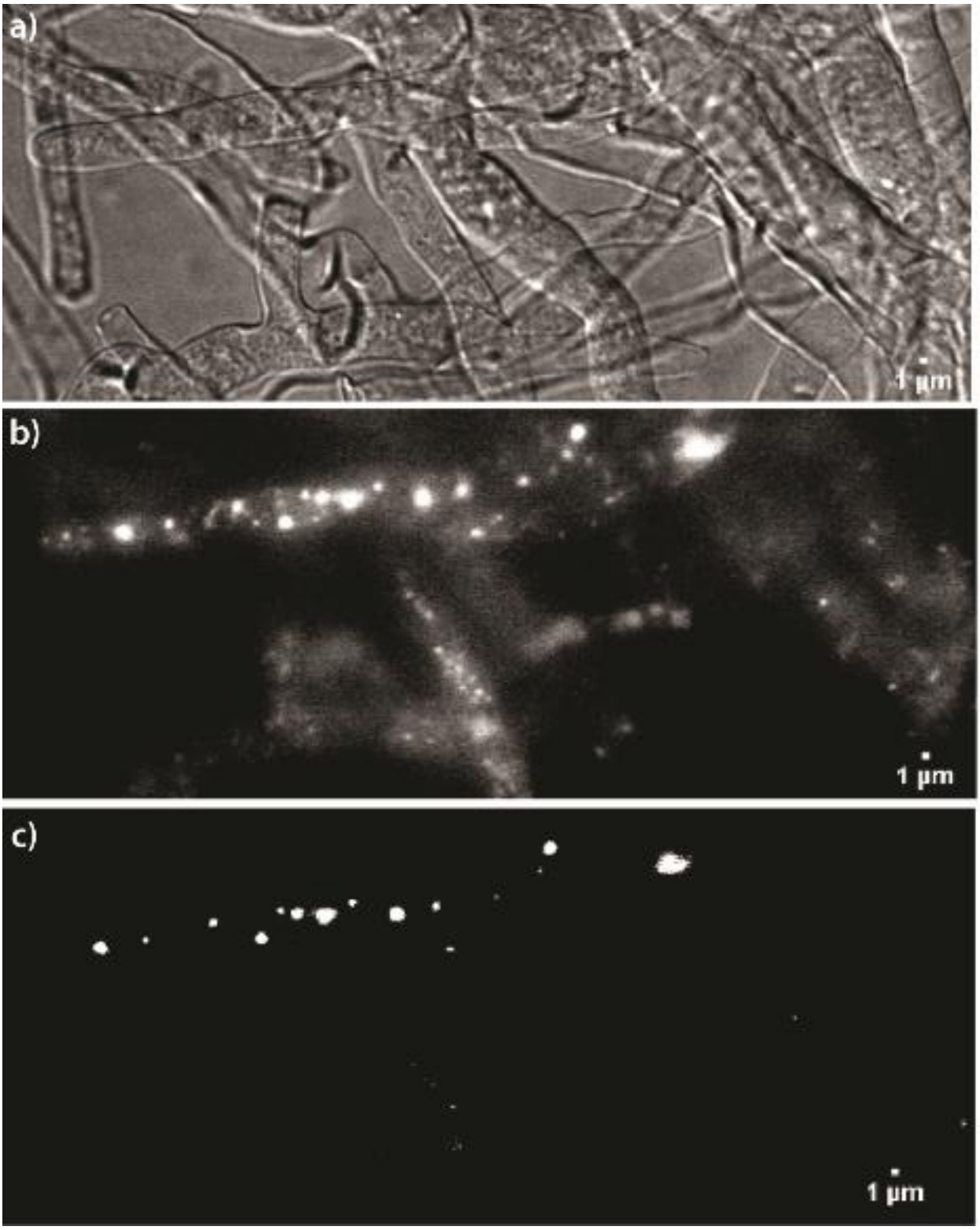

Figure 4.5.3 – Fluorescent microscopy for nuclei counting and quantitation. a) Bright field image identifying cultivar hypha stained with DAPI; b) fluorescent microscopy of same hypha showing stained nuclei; c) processed image to remove noise from auto fluorescence.

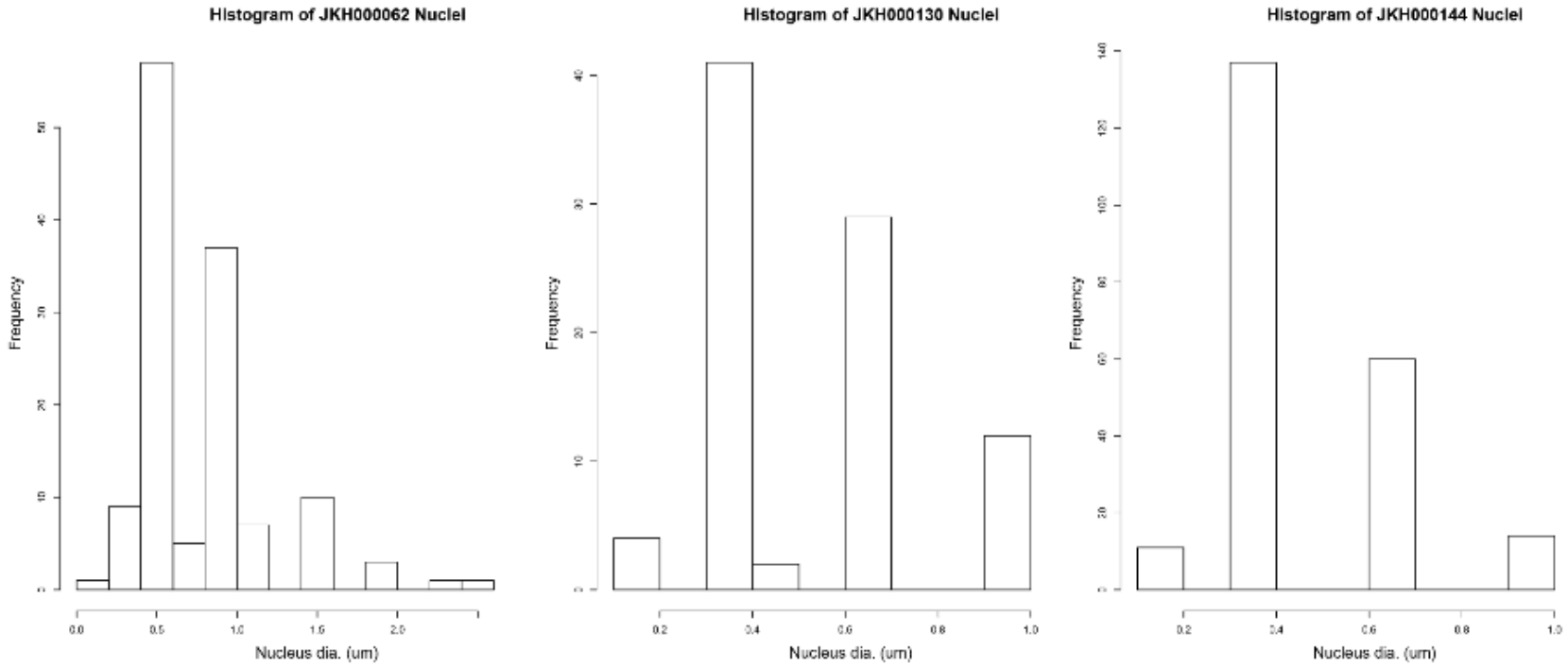

Figure 4.5.4 – Size distribution of nuclei within three different *T. septentrionalis* cultivar isolates. Plots show mean, first and third quartiles, and maximum/ minimum values. Diameter of nucleus is measured in µm.

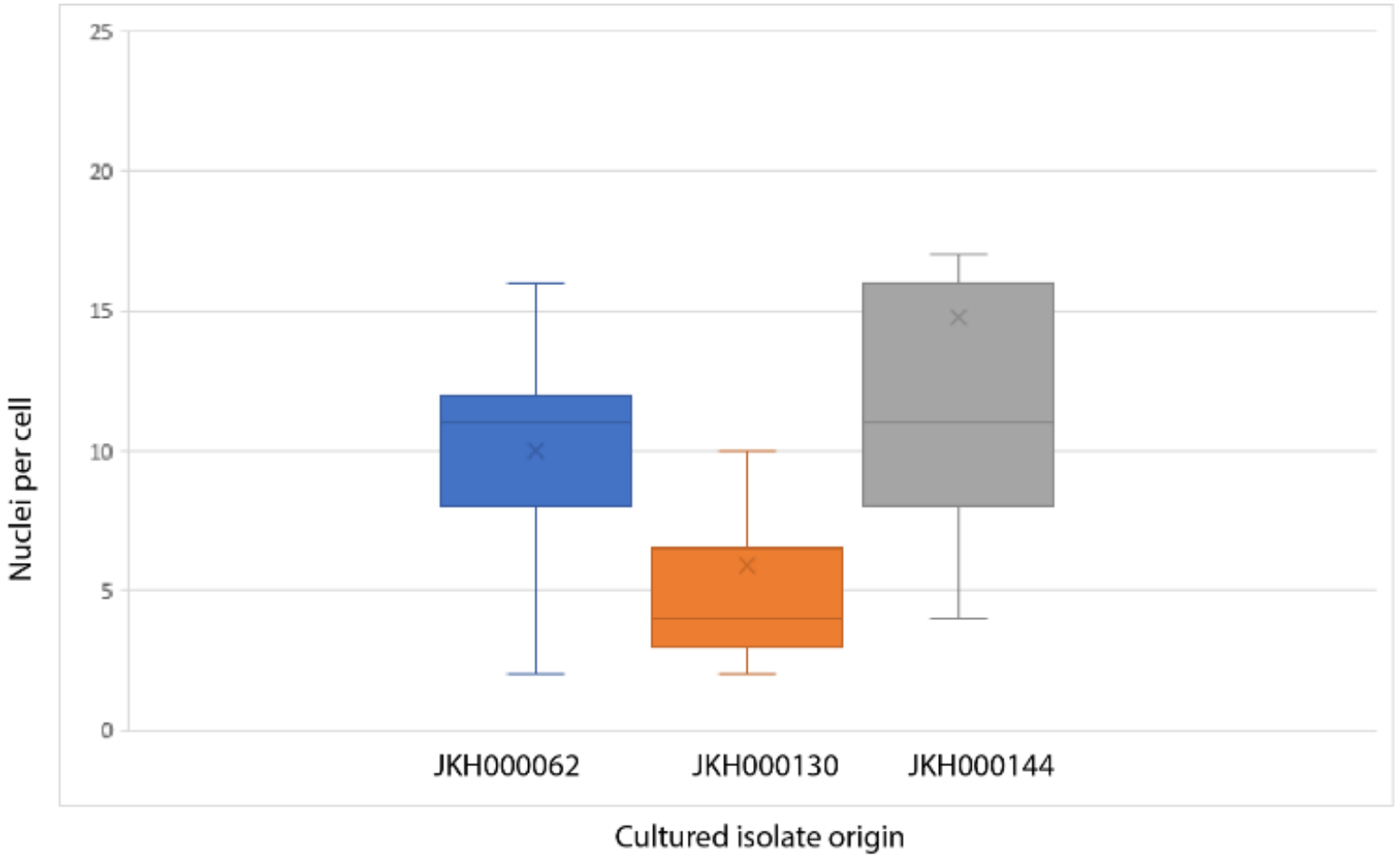

Figure 4.5.5 – Number of nuclei per cell for three different *T. septentrionalis* cultivar isolates. Box and whisker plot of nuclei in 15 cells for each cultured isolate. Plots show mean, first and third quartiles, and maximum/ minimum values. The colony ID each isolate was cultured from is labeled below each plot.

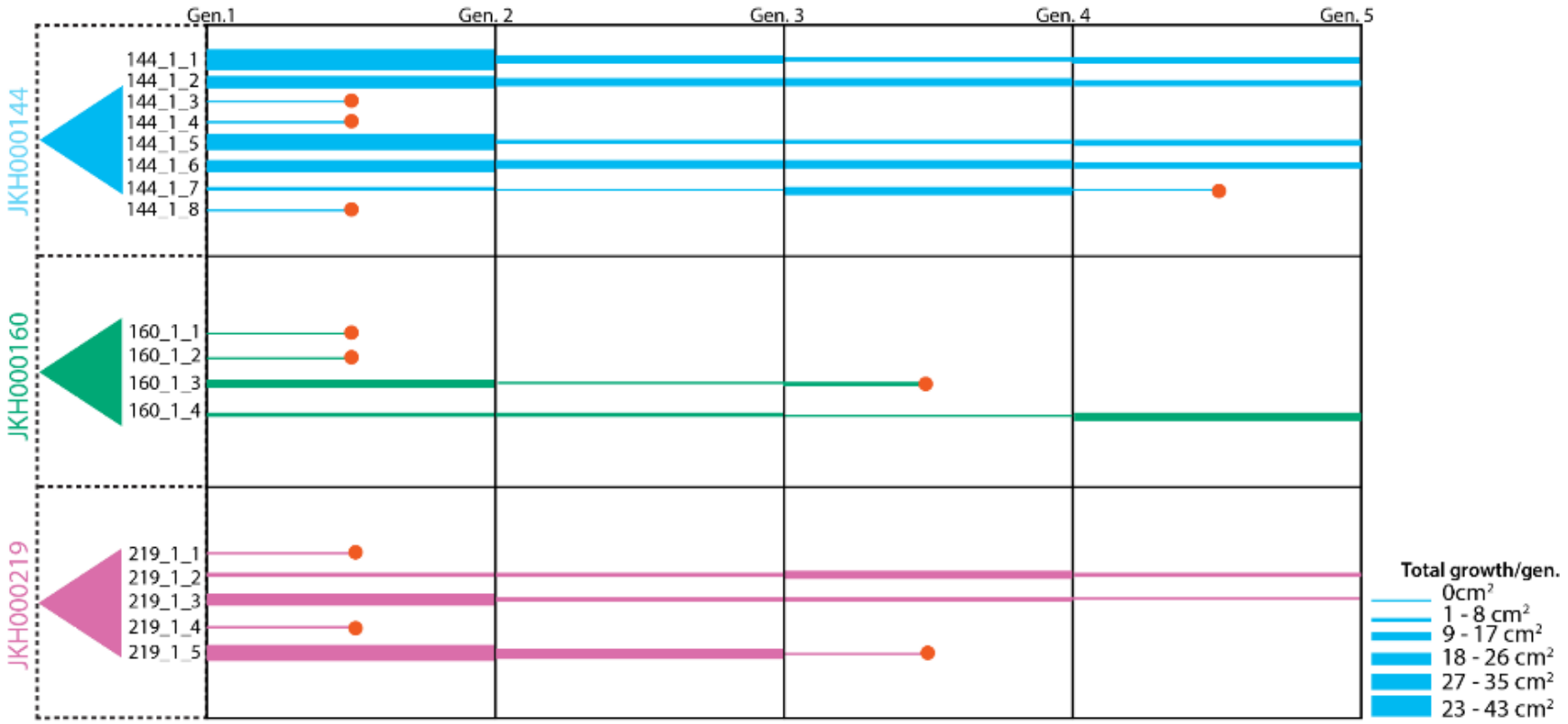

Figure 4.6.1 – Genetic stability experiment genealogy. Genealogical tree of the seventeen lineages of isolates cultured from lab-reared fungus gardens. Colored "JKH" names on the far left indicate from which *T. septentrionalis* host colony number each isolate lineage was derived. Vertical lines divide each subsample generation. Total growth area of each culture is represented by the thickness of the line in each generation. Orange dots represent contaminated subcultures that were removed from the experiment.

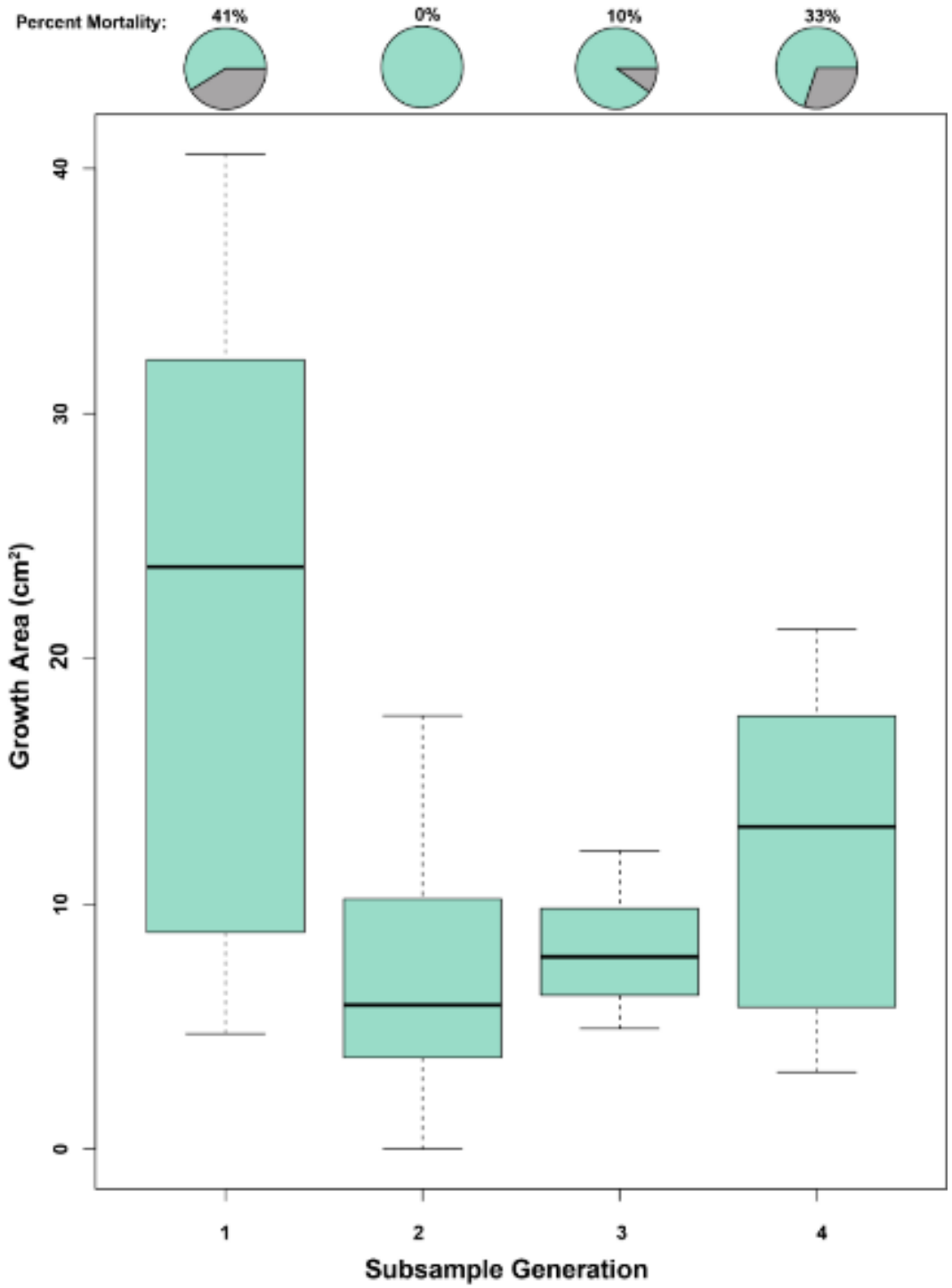

Figure 4.6.2 – Summarized growth rate per subsample generation. Box and whisker plot of aggregated growth rate in each subsampled generation. Plots show mean, first and third quartiles, and maximum/ minimum values. Growth of contaminated samples were excluded, represented instead by the percent mortality pies at the top of the figure; where percent mortality for each generation is depicted in gray and reported above the pies.

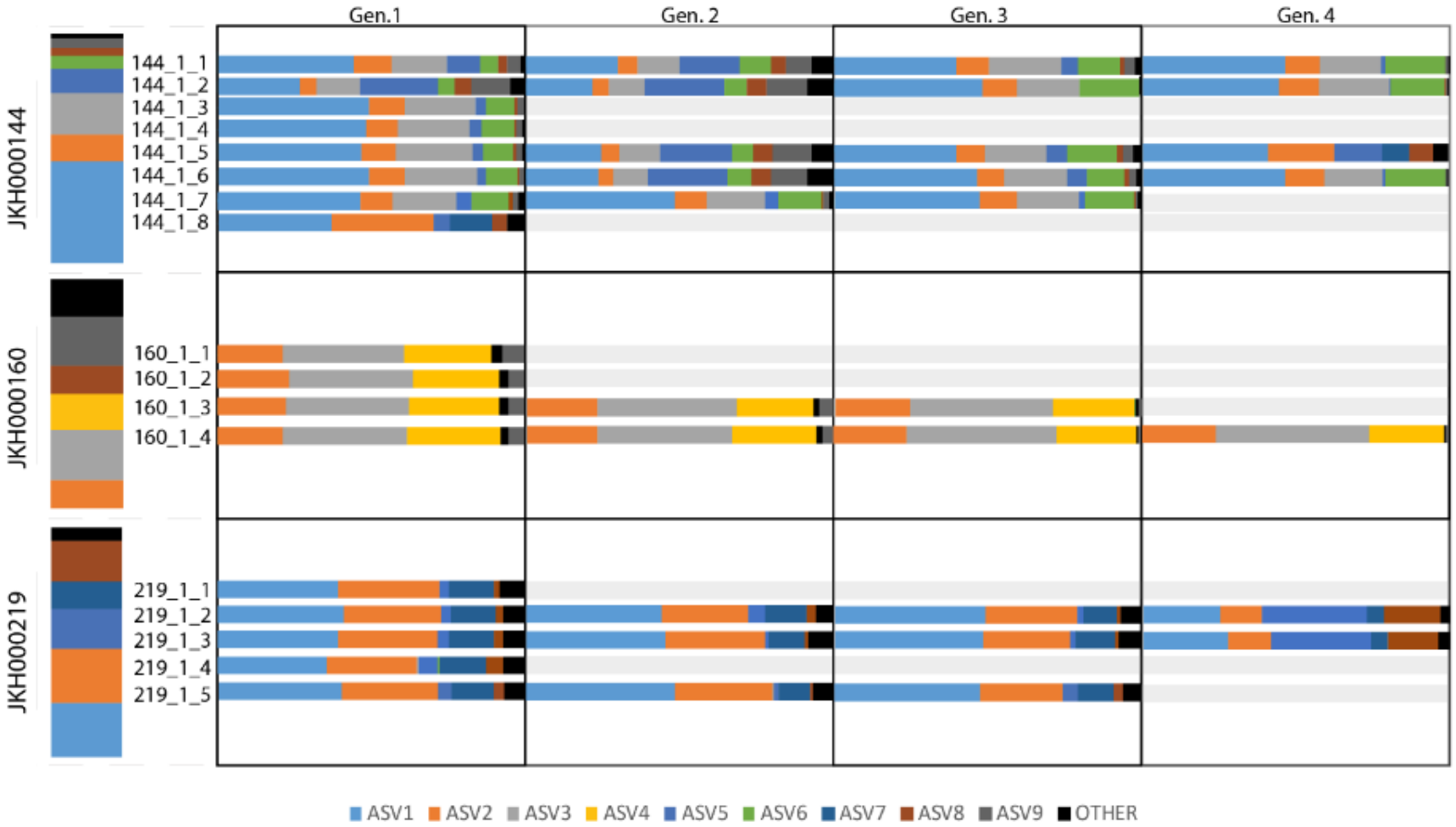

Figure 4.6.3 – ASV abundance genealogy. Genealogical tree of the seventeen lineages of isolates cultured from lab-reared fungus gardens. Colored bars represent the 9 most abundant ASVs and the remaining grouped as "other." Colored "JKH" names on the far left indicate from which *T. septentrionalis* host colony number each isolate lineage was derived. Vertical lines divide each subsample generation. Transition to full light grey bars indicate the contamination of a subculture and its removal from the experiment.

7.2 – Code for analyses

```
##############################################
##File S1 - Code for Bioinformatic Analyses##
##    This is NOT an executable file       ##
##############################################


############################
## DADA2 PROCESSING SEQS ##
############################


#Load packages
library(dada2); packageVersion("dada2")


        #Set path to file containing fastq files
path <- "~/Miseq_datas/MA_LD_DADA2/Fungal\ samples/\ # CHANGE ME to the directory containing the fastq files after unzipping.
list.files(path)


        #Read in names of fastq files
# Forward and reverse fastq filenames have format: SAMPLENAME_R1_001.fastq and SAMPLENAME_R2_001.fastq
fnFs <- sort(list.files(path, pattern="_R1_001.fastq", full.names = TRUE))
fnRs <- sort(list.files(path, pattern="_R2_001.fastq", full.names = TRUE))


# Extract sample names, assuming filenames have format: SAMPLENAME_XXX.fastq
sample.names <- sapply(strsplit(basename(fnFs), "_"), `[`, 1)


        #Plot quality profiles
Qual_prof_Fnfs <- plotQualityProfile(fnFs[1:2])
Qual_prof_Fnrs <- plotQualityProfile(fnRs[1:2])
```

```
        #Filter and trim
# Place filtered files in filtered/ subdirectory
filtFs <- file.path(path, "filtered", paste0(sample.names, "_F_filt.fastq.gz"))
filtRs <- file.path(path, "filtered", paste0(sample.names, "_R_filt.fastq.gz"))
out <- filterAndTrim(fnFs, filtFs, fnRs, filtRs, truncLen=c(240,160),
          maxN=0, maxEE=c(2,2), truncQ=2, rm.phix=TRUE,
          compress=TRUE, multithread=TRUE) # On Windows set multithread=FALSE
head(out)

        #Learn error rates; plot visual for sanity check
errF <- learnErrors(filtFs, multithread=TRUE)
errR <- learnErrors(filtRs, multithread=TRUE)
plotErrors(errF, nominalQ=TRUE)

        #Dereplicate sequences
derepFs <- derepFastq(filtFs, verbose=TRUE)
derepRs <- derepFastq(filtRs, verbose=TRUE)
# Name the derep-class objects by the sample names
names(derepFs) <- sample.names
names(derepRs) <- sample.names

        #Infer samples
dadaFs <- dada(derepFs, err=errF, multithread=TRUE)
dadaRs <- dada(derepRs, err=errR, multithread=TRUE)
dadaFs[[1]] #inspect the data-class object

        #Merge paired reads
mergers <- mergePairs(dadaFs, derepFs, dadaRs, derepRs, verbose=TRUE)
# Inspect the merger data.frame from the first sample
head(mergers[[1]])
```

```
#Construct sequence table
seqtab <- makeSequenceTable(mergers)
dim(seqtab)

#Remove chimeras
seqtab.nochim <- removeBimeraDenovo(seqtab, method="consensus", multithread=TRUE, verbose=TRUE)
dim(seqtab.nochim)
sum(seqtab.nochim)/sum(seqtab) #Inspect drop-off rate as sanity check

#Track reads through the pipepline
getN <- function(x) sum(getUniques(x))
track <- cbind(out, sapply(dadaFs, getN), sapply(dadaRs, getN), sapply(mergers, getN), rowSums(seqtab.nochim))
# If processing a single sample, remove the sapply calls: e.g. replace sapply(dadaFs, getN) with getN(dadaFs)
colnames(track) <- c("input", "filtered", "denoisedF", "denoisedR", "merged", "nonchim")
rownames(track) <- sample.names
head(track)

#Assign taxonomy with unite database -- edit to location of fasta files
taxa <- assignTaxonomy(seqtab.nochim, "~/Miseq_datas/MA_LD_DADA2/Taxa/sh_general_release_dynamic_01.12.2017.fasta", multithread=TRUE)



##########################
## HANDOFF TO PHYLOSEQ ##
##########################

#Load packages
library("phyloseq"); packageVersion("phyloseq")
library("ggplot2"); packageVersion("ggplot2")
```

```
#>>>Question : How do the ASV compositions between raw gardens and cultivar isolates compare?<<<#


        #Read csv tables into R

MA_Taxa_raw <- read.csv(file="C:\\Users\\leede\\OneDrive\\Desktop\\Thesis Materi\\ASV_data\\ASV_data_sheets\\MA_Taxa_raw.csv", header=TRUE, sep=",")

MA_ASVs_raw <- read.csv(file="C:\\Users\\leede\\OneDrive\\Desktop\\Thesis Materi\\ASV_data\\ASV_data_sheets\\MA_Raw_garden_ASVs_12-13.csv", header=TRUE, sep=",")

MA_ASVs_isolates <- read.csv(file="C:\\Users\\leede\\OneDrive\\Desktop\\Thesis Materi\\ASV_data\\ASV_data_sheets\\MA_isolates_ASVs_12-13.csv", header=TRUE, sep=",")


#turn ASV/taxa table .csv into phyloseq object

ps.raw_all <- phyloseq(otu_table(MA_ASVs_raw, taxa_are_rows=TRUE), tax_table(MA_Taxa_raw))

ps.iso_all <- phyloseq(otu_table(MA_ASVs_isolates, taxa_are_rows=TRUE), tax_table(MA_Taxa_raw))

ps.raw_cult <-

ps.iso_cotuult <-


        ###Decontam###

library(phyloseq); packageVersion("phyloseq")

library(ggplot2); packageVersion("ggplot2")

library(decontam); packageVersion("decontam")


        #Import datasets (Sample-ASV and meta table)

meta <- read.csv("seqtable_ASV_ev_raw_meta.csv", header = TRUE, row.names = 1, sep = ",")


        #build phyloseq object

 ps0 <- phyloseq(tax_table(taxa), otu_table(seqtab.nochim, taxa_are_rows = FALSE))


        #assign meta data to sample_data

sample_data(ps0) <- meta
```

```
> sample_data(ps0)$is.neg <- NA
> sample_data(ps0)$Sample_or_Control <- NA

sample_data(ps0)$is.neg <- sample_data(ps0)$Sample_or_Control == "Control Sample"
contamdf.prev <- isContaminant(ps0, method="prevalence", neg="is.neg")
table(contamdf.prev$contaminant)
less

contamdf.freq <- isContaminant(ps0, method="frequency", conc="quant_reading")
head(contamdf.freq)



        #rarefy to equal depth (1000 reads)
 ps_rar_1000 <- rarefy_even_depth(ps, sample.size = 1000)

######Aplha-diversity##############

        #plot bars for ASV composition, fill w/ family
pdf("MA_bars_raw_rar")
plot_bar(raw_rar, fill="Genus")
dev.off()

        #plot alpha div. (Chao1 & Shannon)plot
pdf("MA_adiv_raw_rar.pdf")
plot_richness(raw_rar, measures=c("Chao1", "Shannon"))
dev.off()
```

#generate figure from ps object

ps.cult <- phyloseq(otu_table(ASV_table_cultivar_ed, taxa_are_rows=FALSE), tax_table(taxa_table_cultivar_ed))

#general parameters for tree building

raxmlHPC -m GTRGAMMAI -n raw_garden_cult_ASVs_1213 -s msa_seqs.fasta -p 9243 -f a -x 1112 -N 500 -T 24

#BLASTn input: make directory and BLAST against it, output in readable table

makeblastdb -in reference_seqs.fasta -dbtype nucl

blastn -query trimmed_msa.ASV_ev_ref.fasta -db reference_seqs.fasta -evalue 1e-5 -out ASV_Ev_cult_BLAST.blastn -outfmt 7

#make csv file from ps object

write.csv(otu_table(ps_rar_1000), file = "rarified_1K_ASVtab.csv")

####################

#SparCC correlation#

####################

python SparCC.py Rar_1k_ASVtab.txt -i 10 -i 5 --cor_file=outputs/basis_corr/cor_sparcc.out -a sparcc

python SparCC.py Rar_1k_ASVtab.txt -i 10 -i 5 --cor_file=outputs/basis_corr/cor_pearson.out -a pearson

python SparCC.py Rar_1k_ASVtab.txt -i 10 -i 5 --cor_file=outputs/basis_corr/cor_spearman.out -a spearman

python MakeBootstraps.py outputs/Rar_1k_ASVtab.txt -n 100 -t permutation_#.txt -p outputs/pvals/

python SparCC.py example/pvals/permutation_0.txt -i 5 --cor_file=outputs/pvals/perm_cor_0.txt

python SparCC.py example/pvals/permutation_1.txt -i 5 --cor_file=outputs/pvals/perm_cor_1.txt

python SparCC.py example/pvals/permutation_2.txt -i 5 --cor_file=outputs/pvals/perm_cor_2.txt

python SparCC.py example/pvals/permutation_3.txt -i 5 --cor_file=outputs/pvals/perm_cor_3.txt

python SparCC.py example/pvals/permutation_4.txt -i 5 --cor_file=outputs/pvals/perm_cor_4.txt

python PseudoPvals.py example/basis_corr/cor_sparcc.out outputs/pvals/perm_cor_#.txt 5 -o example/pvals/pvals.one_sided.txt -t one_sided

python PseudoPvals.py example/basis_corr/cor_sparcc.out outputs/pvals/perm_cor_#.txt 5 -o example/pvals/pvals.one_sided.txt -t two_sided

#run shell script to resample datasets:

#!/bin/bash

##########################

#

Simple SparCC wrapper

#

Use:

chmod 755 sparccWrapper.sh

./sparccWrapper.sh &

#

Author: Karoline Faust

#

##########################

here goes the path to the input file

INPUT_PATH="Rar_1k_ASVtab.txt"

here goes the path to the output folder

OUTPUT_PATH="ResultsSparCC"

here goes the sparcc root folder

SPARCC_PATH="sparcc"

ITER=10

BOOT_ITER=100

```
WORKDIR=$PWD


mkdir $OUTPUT_PATH/Resamplings

mkdir $OUTPUT_PATH/Bootstraps


cd $SPARCC_PATH


python SparCC.py $INPUT_PATH -i $ITER --cor_file=$OUTPUT_PATH/Rar_1k_ASVtab.txt >
$OUTPUT_PATH/sparcc.log


python MakeBootstraps.py $INPUT_PATH -n $BOOT_ITER -o
$OUTPUT_PATH/Resamplings/boot


# compute sparcc on resampled (with replacement) datasets

for i in 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61
62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91
92 93 94 95 96 97 98 99

do

    python SparCC.py $OUTPUT_PATH/Resamplings/boot_$i.txt -i $ITER --
cor_file=$OUTPUT_PATH/Bootstraps/sim_cor_$i.txt >> $OUTPUT_PATH/sparcc.log

done


# compute p-value from bootstraps

python PseudoPvals.py $OUTPUT_PATH/Rar_1k_ASVtab.txt
$OUTPUT_PATH/Bootstraps/sim_cor $ITER -o $OUTPUT_PATH/pvals_two_sided.txt -t
'two_sided'  >> $OUTPUT_PATH/sparcc.log


# visualization requires parsing and thresholding the p-value OTU matrix

-------------------------------------

chmod 755 sparccWrapper.sh

./sparccWrapper.sh &


        #create heatmap from correlation matrix (in R)
```

```
library(RColorBrewer)

library(gplots)

data = read.table(file = "cor_spearman.out", sep = '\t', header = TRUE)

matrix <- as.matrix(data[,-1])

breaks=seq(0,1,0.01)

colors=rev(brewer.pal(11,"Spectral"))

colors=colorRampPalette(colors)(100)

pdf("SparCC_heatmap_rar.pdf")

heatmap.2(matrix, reorderfun=function(d,w) reorder(d, w, agglo.FUN=mean), distfun=function(x) dist(x, method="euclidean"), hclustfun=function(x) hclust(x, method="ward.D"), col=colors, density.info="none", trace="none", breaks=breaks, cexRow=0.1, cexCol=0.1)

dev.off()
```